

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:32:48 ; Search time 31.47 Seconds  
(without alignments)  
854.510 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012  
Sequence: 1 MILGILASLVNGACLPFLMPV.....QELLNRNDYFKLVNQSVM 1195

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3543	58.9	1280	2	US-08-583-276-19
2	3543	58.9	1280	6	5206352-4
3	3528.5	58.7	1279	2	US-08-784-649A-2
4	3527	58.7	1280	2	US-08-752-447-2
5	3517.5	58.5	1275	3	US-09-120-513-2
6	3517.5	58.5	1275	4	US-09-450-105-2
7	2282	38.0	1334	4	US-08-996-545-2
8	2282	38.0	1334	4	US-09-328-320-2
9	2273.5	37.8	1349	2	US-08-612-734B-2
10	2166	36.0	1307	1	US-08-395-246C-2
11	2116	35.2	1408	1	US-08-612-521-2
12	1695.5	28.2	1308	3	US-08-996-644-2
13	1695.5	28.2	1308	3	US-09-352-552-2
14	1664	27.7	1302	1	US-08-233-537-2
15	962.5	16.0	327	1	US-08-463-092B-9
16	962.5	16.0	327	2	US-08-460-907B-9
17	761.5	12.7	791	1	US-08-394-880B-2
18	692	11.5	686	4	US-09-061-764A-15
19	687.5	11.4	1528	1	US-08-463-092B-6
20	687.5	11.4	1528	1	US-08-462-109A-6
21	687.5	11.4	1528	2	US-08-460-907B-6
22	687.5	11.4	1528	3	US-08-463-179A-6
23	687.5	11.4	1528	3	US-08-461-384B-6
24	671.5	11.2	1531	1	US-08-463-092B-4
25	671.5	11.2	1531	1	US-08-462-109A-4
26	671.5	11.2	1531	2	US-08-460-907B-4
27	671.5	11.2	1531	3	US-08-463-179A-4

28	671.5	11.2	1531	3	US-08-461-384B-4	Sequence 4, Appl
29	671	11.2	748	4	US-09-061-764A-19	Sequence 19, Appl
30	669	11.1	1437	3	US-09-061-7400-2	Sequence 2, Appl
31	669	11.1	1453	2	US-09-001-273-2	Sequence 2, Appl
32	669	11.1	1453	4	US-08-843-459A-2	Sequence 2, Appl
33	668.5	11.1	1531	1	US-08-141-893-2	Sequence 2, Appl
34	668.5	11.1	1531	1	US-08-463-092B-2	Sequence 2, Appl
35	668.5	11.1	1531	2	US-08-462-109A-2	Sequence 2, Appl
36	668.5	11.1	1531	2	US-08-460-907B-2	Sequence 2, Appl
37	668.5	11.1	1531	3	US-08-463-179A-2	Sequence 2, Appl
38	668.5	11.1	1531	3	US-08-461-384B-2	Sequence 2, Appl
39	668.5	11.1	1531	3	US-08-407-207A-2	Sequence 2, Appl
40	648	10.8	707	3	US-08-772-270A-4	Sequence 4, Appl
41	629	10.5	653	4	US-09-061-764A-2	Sequence 2, Appl
42	628.5	10.5	1621	4	US-08-972-927-3	Sequence 3, Appl
43	626	10.4	1622	4	US-08-972-927-6	Sequence 6, Appl
44	620	10.3	711	3	US-08-772-270A-12	Sequence 12, Appl
45	599.5	10.0	694	2	US-08-895-522-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-583-276-19  
; Sequence 19, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:  
; APPLICANT: McDonagh, Kevin T.  
; APPLICANT: Nienhuis, Arthur  
; APPLICANT: Tolstoshev, Paul  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Giffalian,  
; ADDRESSEE: Cecchi & Stewart  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: DM4.V2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,276  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/332,444  
; FILING DATE: 31-OCT-1994  
; APPLICATION NUMBER: 07/887,712  
; FILING DATE: 22-MAY-1992  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1280 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; DESCRIPTION: protein  
; US-08-583-276-19  
Query Match 58.9%; Score 3543; DB 2; Length 1280;  
Best Local Similarity 55.5%; Pred. No. 0;  
Matches 681; Conservative 245; Mismatches 265; Indels 36; Gaps 9;  
QY 1 MILGILASLVNGACLPFLVIGEMSD-----NLISGLVOT--NTYSF----- 43

```

Db 51 MYVGTAAIHHGAGLPLMLVGEWTDIFANAGNLEDMSNTNSDNDGFFPNLEED 110
Oy 44 --RLTYVVGIGVALALFYGIQISLMTTAAQRTKIRKOPFHSHVLAODIGFSDCDGE 101
Db 111 MTRVAYVSGIGAGVLAAYTIOVSFWCLAGRQIHKIRKOPFHSHVLAODIGFSDCDGE 170
Oy 102 LNTKMT-DIDRISDIDGKIALLPQNMSTFSGIAGLVGKRLVLTSLSPILMASAA 160
Db 171 LNTRLTDVSKINEYIGDKIGFQSMATFTGFIYGTGKMLVLTSLSPILMASAA 230
Oy 161 ACSRWVLSLTSKELSAISKAGAVAEVLSSTRTVIAFPAQEKELORYONLKDADKFGIK 220
Db 221 VWAKLTSFTDKELAYAKAGAVAEVLAARTVIAFGOKKELERYKNLLEAKRIGIK 290
Oy 221 RTIASVLSIAGVYFPMNGTGYLAFWYGTSLNGBPGYTIGVLAVFVSVHSSYCIGA 280
Db 291 KAITANISIGAFLIIVASYALAFWYGTTLVLSGE--YSIGVLTVPFVSILGAFSVGOA 348
Oy 281 VPHETFAIANGAAPHIFOVIDKKSINDNESTAGYKPSIEGTFEKNVSEFVSPSPSIK 340
Db 349 SPSEIAFANANGAAVEIKIIDNKPSIDSYSGHKPDNIKONLEFRVHVSYPSPREKVK 408
Oy 341 ILKGLNIRKSGEYVALVGLNGSGKSTVVOQLRLYDPDDGFIWVDENDIRALNVRYHRD 400
Db 409 ILKGLNIRKSGEYVALVGLNGSGKSTVVOQLRLYDPDDGFIWVDENDIRALNVRYHRD 468
Oy 401 HIGVYSEPVLEFGTTISNNIKYGRDVTDEMERARARANAYDFMEPPKNFTLVGSKG 460
Db 469 IIGVYSEPVLEFATTIAENIRYGRONVTMDETERAKVKNANAYDFMKLPKFDLTVGSKG 528
Oy 461 AOMSGGOKRIATAALVARNPKIILLDEATSAJDSEKSAVOALEKASKGRTTIVVAHR 520
Db 529 AOLSGGOKRIATAALVARNPKIILLDEATSAJDSEKSAVOALDKARKGRTTIVVAHR 588
Oy 521 LSTRSADLIYTLKDGMLAEKGAHAELMAKRGLYSLVMSQ-----DIKADQEMESMT 574
Db 589 LSTYANAVIAGFDGVTVEKGNHDELMKEKGIYFKLVMTQAGNEVELENADSEKSEI 648
Oy 575 YSTERKTSPLHSHV-----KSIK-SDFIDKAEESTQSKETSLPEVSLKTIKINKREW 627
Db 649 DALESNDSSSLIRKSTRSVAGSOADPKSLTKALDLSIPVFWIRMKNTLW 708
Oy 628 PEVVLGTASVYNGTVHVPFSIIFAKIITMFGN-NDKTTLKHAEIYEMIEFVLGICFV 686
Db 709 PYFVGVCAIINGSLQAFAPAIIFSKIIGVTRIDPEPKQNSLFLPLALGIIISFI 768
Oy 687 SYFMGELFYGRAGETLIRLRLHAFKAMLYODIANFDEKENSTGGLTTTLAIDAOIGA 746
Db 769 TFFLOGFPEFGAGETLIRLRLYMFERSMLRODVSWFDDPKNTTGALTTRLANDAQVKG 828
Oy 747 TGSRTGVLTONATNGLSVITISFYGMETPLIISIAVLAVTGMEIAAMFANGKOKO 806
Db 829 IGSRLATVTONANLGTIIISFYGMOTLLLAIVPLIAGVEKMMISGOLAKOK 888
Oy 807 ELKHAGKATATELENIRIVSLTREKAFEOYMEEMLOTOHNTSKAOIIGSCYAFSHAF 866
Db 889 ELEAGKATATELENIRIVSLTREKAFEOYMEEMLOTOHNTSKAOIIGSCYAFSHAF 948
Oy 867 IYFAYAGFRGAYLIQAGKMTPEGMFLVFETALAYGAMAGTGLVLADEYSKAGGAHL 926
Db 949 MYFSYAGFRGAYLIQAGKMTPEGMFLVFETALAYGAMAGTGLVLADEYSKAGGAHL 1008
Oy 927 FALLEKKNIRSRQEGKKPDTCEGNLEFRMSFPYPCRPDYFLIRGLSLSTERKTYAF 986
Db 1009 IMTLEKTPLDISYSTEGLMPNTLBSNVFGEVFNYPTRPDIPLYOGSLVLEKQOTLAL 1068
Oy 987 VGSSECGKSTVOLRLYDPVOGVLDPGVDAKELANQWLSQAIYPOEFVLENCISA 1046
Db 1069 VGSSECGKSTVOLRLYDPVOGVLDPGVDAKELANQWLSQAIYPOEFVLENCISA 1128
Oy 1047 ENIAYGNSRVVPLDEIKEAANANHSFIEGLPERKYNTOVGLKAQOLSGGOKRIATAAR 1106

```

```

Db 1129 ENIAYGNSRVVSEIIVRAKEANIHAFIESLPKRYSTKVGDKGTQSGGOKRIATAAR 1188
Oy 1107 ALLOKPKIILLDEATSAJDSEKSAVOALDKARKGRTTIVVAHR 1166
Db 1169 ALVROPHILLDEATSAJDSEKSAVOALDKARKGRTTIVVAHR 1248
Oy 1167 GKIKGQTHQELLNRNDIYFKLVNAOS 1193
Db 1249 GRVKEGTHQELLNRNDIYFKLVNAOS 1275

RESULT 2
5206352-4
Patent No. 5206352
Applicant: Roninson, Igor B.; Pastan, Ira H.; Gottesman, Michael M.
TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/622,836
FILING DATE: 24-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 892,575
FILING DATE: 01-AUG-1986
APPLICATION NUMBER: 845,610
FILING DATE: 28-MAR-1986
SEQ ID NO: 4
LENGTH: 1280
5206352-4

Query Match 58.9%; Score 3543; DB 6; Length 1280;
Best Local Similarity 55.5%; Pred. No. 0;
Matches 681; Conservative 245; Mismatches 265; Indels 36; Gaps 9;

Oy 1 MILGILASLVNACDPLMPLVLGEMSD-----NLISGLVOT--NTYSFF----- 43
Db 51 MYVGTAAIHHGAGLPLMLVGEWTDIFANAGNLEDMSNTNSDNDGFFPNLEED 110
Oy 44 --RLTYVVGIGVALALFYGIQISLMTTAAQRTKIRKOPFHSHVLAODIGFSDCDGE 101
Db 111 MTRVAYVSGIGAGVLAAYTIOVSFWCLAGRQIHKIRKOPFHSHVLAODIGFSDCDGE 170
Oy 102 LNTKMT-DIDRISDIDGKIALLPQNMSTFSGIAGLVGKRLVLTSLSPILMASAA 160
Db 171 LNTRLTDVSKINEYIGDKIGFQSMATFTGFIYGTGKMLVLTSLSPILMASAA 230
Oy 161 ACSRWVLSLTSKELSAISKAGAVAEVLSSTRTVIAFPAQEKELORYONLKDADKFGIK 220
Db 221 VWAKLTSFTDKELAYAKAGAVAEVLAARTVIAFGOKKELERYKNLLEAKRIGIK 290
Oy 221 RTIASVLSIAGVYFPMNGTGYLAFWYGTSLNGBPGYTIGVLAVFVSVHSSYCIGA 280
Db 291 KAITANISIGAFLIIVASYALAFWYGTTLVLSGE--YSIGVLTVPFVSILGAFSVGOA 348
Oy 281 VPHETFAIANGAAPHIFOVIDKKSINDNESTAGYKPSIEGTFEKNVSEFVSPSPSIK 340
Db 349 SPSEIAFANANGAAVEIKIIDNKPSIDSYSGHKPDNIKONLEFRVHVSYPSPREKVK 408
Oy 341 ILKGLNIRKSGEYVALVGLNGSGKSTVVOQLRLYDPDDGFIWVDENDIRALNVRYHRD 400
Db 409 ILKGLNIRKSGEYVALVGLNGSGKSTVVOQLRLYDPDDGFIWVDENDIRALNVRYHRD 468
Oy 401 HIGVYSEPVLEFGTTISNNIKYGRDVTDEMERARARANAYDFMEPPKNFTLVGSKG 460
Db 469 IIGVYSEPVLEFATTIAENIRYGRONVTMDETERAKVKNANAYDFMKLPKFDLTVGSKG 528
Oy 461 AOMSGGOKRIATAALVARNPKIILLDEATSAJDSEKSAVOALEKASKGRTTIVVAHR 520
Db 529 AOLSGGOKRIATAALVARNPKIILLDEATSAJDSEKSAVOALDKARKGRTTIVVAHR 588
Oy 521 LSTRSADLIYTLKDGMLAEKGAHAELMAKRGLYSLVMSQ-----DIKADQEMESMT 574

```



Db 589 LSTVRNADVIAGFDGVIYVEKGNHDELMKEGIFKLVMTQTAGNEVLENAADESKSEI 648  
QY 575 YSTERKNSLPLHSV-----KSIR-SDFIDKAESTOSKEISLPEVSLKTLKLNKPEW 627  
Db 649 DALEMSNDSSRSSLRKSTRSRVSGSOADRKLSTKEALDESIPVPSFMRKLNLEW 708  
QY 628 PFWVLGLASLVNCTVHPVPSIIFAKIITMGCN-NDKTTLKHDAEISMITVILGVICFV 686  
Db 709 PYFVGVFCALINGGLDPAFAIIFRSKIIGVTRIDDPETKRONSLFSLFLALGIISFI 768  
QY 687 SYFMQGLFYGRAGEILTMRLRLHLAFKAMLYODIAMFDEKENSTGGLTTLAIDAOIOGA 746  
Db 769 TFFLOGFTFGAGELITKRLRYMFRSLRQDVSNFDDPKNTTGALTTRLANDAQAQVGA 828  
QY 747 TGSRIQVLTQNTMNGLSVLIISFTYGMETPLILISAPVLAVTGMIETAMTGFANKDKQ 806  
Db 829 IGSRLAVTQNTMNGIISFTYGMQTLTLLAIVPIIILAGVEMKMLSGCALKDKK 888  
QY 807 ELKAGKATATPALTNIIRIYVSLTREKAFEQMYEEMLOTOHNTSKKAOIISCFASHAF 866  
Db 889 ELEAGKATEAIENFRIVSLTQDKREHMYAOSLOVPYHNSLRKAHIFGITSFTQAM 948  
QY 867 IYFVYAGFRGAYLIOGRMTPGEMFIYFAIYAGAMAIKTLVLAPYKAKSGAHL 926  
Db 949 MYESTAGCFRGALIVAHKMLSFEDVILVFSAYVFGAMAVQVSSFADYAKAKISAH 1008  
QY 927 FALLEKPNIDSRSQEGKPDTCGNTFEFVSEFFPCRPDVLITRGLSISIERGKTVAF 986  
Db 1009 IMIEKTPILDISYESTEGMPTLEGVNFVGFVNYPRPDIPLYGLSTLEKKGQOTLAL 1068  
QY 987 VGSSGCKSTSVOLLRHYDVQGOVLFDGDAKELNQMRLSOATVPOBPVLFNCISA 1046  
Db 1069 VGSSGCKSTSVOLLRHYDVQGOVLFDGDAKELNQMRLSOATVPOBPVLFNCISA 1128  
QY 1047 ENIYAGDSRVYPLDEIKKAAANANHSIFGLEPEKYNTQVGLKQAOISGOKORLATA 1106  
Db 1129 ENIYAGDSRVYPLDEIKKAAANANHSIFGLEPEKYNTQVGLKQAOISGOKORLATA 1188  
QY 1107 ALLOPKILLDEATSALENDSEKVOYHALLDKARTGRTCLVYTHRLSAIQNADLVLAN 1166  
Db 1189 ALVQPHILLDEATSALENDSEKVOYHALLDKARTGRTCLVYTHRLSAIQNADLVLAN 1248  
QY 1167 GKIKGQTHOELLNRDITFKLVNAOS 1193  
Db 1249 GRVKEHGHQOLLAQGIYFSMVSOA 1275

RESULT 3  
US-08-784-649A-2  
Sequence 2, Application US/08784649A  
Patent No. 5830697  
GENERAL INFORMATION:  
APPLICANT: SIKIC, Branimir I  
APPLICANT: Chen, Gang  
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,649A  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: Reg.No. 5830697 36,677  
REFERENCE/DOCKET NUMBER: 06037/007001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1279 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-784-649A-2

Query Match 58.7%; Score 3528.5; DB 2; Length 1279;  
Best Local Similarity 55.4%; Pred. No. 0;  
Matches 680; Conservative 246; Mismatches 264; Indels 37; Gaps 10;

QY 1 MILGILASLVNGACPLMPVLYGEMSD-----NLISGCLVQT--NTYSFF----- 43  
Db 51 MYVGTALAIHGAGPLMLVGVEMTDIFANAGNLEDMSTNRSDINDGCFWNLBED 110  
QY 44 --RLTIYVVGIGVALITGCIQISLMTTARQTRIRKOPFHSHVLADDIGFDCDICE 101  
Db 111 MTRVAYYSIGAGVLAAYIOWSFMCILAGROHRIKOPFHALMROEIGFVDHVG 170  
QY 102 LNTFRMT-DIDKIDGIGKIALLPONMSTFSGVLGVKCKLTVLSTPLIMASAA 160  
Db 171 LNTRLTDVSKINEYIGKIGMFPQSMATFTFGFVLTGRGKLLVLIALSPVGLSAA 230  
QY 161 ACSRWYISLTSKELSAYSKAGAAVEVLSTRTVIAFRAQEKELQRYTQNLKANDPFSIK 220  
Db 221 VWAKLTSFTKELLAYAKAGAAVEVLAAIRTVIAFGQKELRYNNKLEEARIGIK 290  
QY 221 RTIASKVSLGAVYFFMNGTYGLAFWYGTSLINGEPGTTIGTVLAVFVSVHSSICGAA 280  
Db 291 KAITANISIGAAFLIIVASYALAFWYGTTLVLSGE--YSIGOVLTV-PSVLGAFSVGOA 347  
QY 281 VPHFETALARGAFHIOVDIKRPSIDNEFSTAGKPSIGETVEFKNVSFVPSRPSIK 340  
Db 348 SPSTAFANARGAAYEIKRIIDNKPISIDYSKSGKHPDITKGNLEFRNHSYPSRKEK 407  
QY 341 ILKGLNRIKSGEYVALVGLNMGSKSTVOLLQRIYDPDGFINVENDIRALNRYRD 400  
Db 408 ILKGLNRIKSGEYVALVGLNMGSKSTVOLLQRIYDPDGFINVENDIRALNRYRD 467  
QY 401 HIGVVSQEPVLEFGTTISNNIKYGRDVTDEMEKARAREANAYDTMEFPNKNNTLVGEGK 460  
Db 468 IIGVVSQEPVLEFATITIAIRKIRYGRNVMTDETEKAVKANAADFTMKLPKHKDTLVGEG 527  
QY 461 AOMSGGOKORTAIARALVRNPKIILLDEATSALEDSASVAALEKASGRTTVIVAR 520  
Db 528 AOLSOGGOKORTAIARALVRNPKIILLDEATSALEDSASVAALEKASGRTTVIVAR 587  
QY 521 LSTISADLVTLKDGMLAEKGAHAELAKRGVLSLVMSQ-----DIRKAQDEQMSWT 574  
Db 588 LSTVRNADVIAGFDGVIYVEKGNHDELMKEGIFKLVMTQTAGNEVLENAADESKSEI 647  
QY 575 YSTERKNSLPLHSV-----KSIR-SDFIDKAESTOSKEISLPEVSLKTLKLNKPEW 627  
Db 648 DALEMSNDSSRSSLRKSTRSRVSGSOADRKLSTKEALDESIPVPSFMRKLNLEW 707  
QY 628 PFWVLGLASLVNCTVHPVPSIIFAKIITMGCN-NDKTTLKHDAEISMITVILGVICFV 686  
Db 708 PYFVGVFCALINGGLDPAFAIIFRSKIIGVTRIDDPETKRONSLFSLFLALGIISFI 767  
QY 747 TGSRIQVLTQNTMNGLSVLIISFTYGMETPLILISAPVLAVTGMIETAMTGFANKDKQ 806  
Db 828 TFFLOGFTFGAGELITKRLRYMFRSLRQDVSNFDDPKNTTGALTTRLANDAQAQVGA 827







```

:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/996,545
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11766
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1334 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-996-545-2

```

```

Query Match      38.0%; Score 2282; DB 2; Length 1334;
Best Local Similarity 41.3%; Pred. No. 1,66-197;
Matches 487; Conservative 225; Mismatches 434; Indels 34; Gaps 9;

```

```

OY 47 LYYGIGVALLFGYIQLSLMTTAAKQKRIKQFPHSVLAQDIGWFDSCDIGELNRM 106
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 153 LYFVLGIGEFYVTVGVFGEHATQKIREYLESILHQNIGYFKLQAGVETTRI 212
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 107 T-DIDKIDGIDKIALLFQNNSTFSLGVLGVGKMLTIVLTSTPLMAAASCSM 165
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 213 TADNVLIDGISEKYLTLTALFTAFITAYVYKMKLALCSSTIYALVLTMGGSQF 272
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 166 VISLSKELSAVSKAGAAVEEVLISRTVIAFRAQEKELQRTOMLKADKDGIRKTAS 225
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 273 IIKYKSKLSDSGAGCTYAAEEVSISSINATAGTODKLAKQVEVLDEKKNGTNOIYM 332
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 226 KVSIGAVYFPNGTYGLAFWGTSLILNGEPYITGVLAFFSVIHSYCIAGAAVPHFE 285
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 333 GFMIAGMGLMANSYGLGFWMGSRFLVDG--AVDVGDLITVLMALIGSFSGVSPNAQ 390
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 286 TFAIRGAFFHFOYIDKKPSIDNSTAGYKPESTIEGTYEFKNVFNPSRSITILKGL 345
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 391 AFTNVAALAAKFFGTIDQSPLDPSNKGKTLDFEGHLELNVRHIIYPSRREVTYMEQV 450
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 346 NLRKSGEVALVGLNGSGKSTVVOQLQRLYDPDGFIVDENDIRALNVRHYRDHIGVY 405
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 451 SLSPMAGKTTALVGPSSGSKSTVGLVERFYMPVKGTYVLLDGHDIKDLNRLRQOISLY 510
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 406 SOEPLVFGTTISNNIKYG-----RDVYDEEMERARERANAYDFIEMPNKFNFTLV 456
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 511 SOEPLVFGTTIYKNIHIGLIGTKYENESDQKVELIENAKMANAHDFITALPEGETNV 570
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 457 GEKGQMGSGGOKORAIARALVNRNKKILIDPATSLDSEKSAVOALAEKSKGRTIVY 516
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 571 GORGELLSGGOKORIAIARAAYSDPKILLDEATSLDTRKSGVVOALAEKAEKRTIVY 630
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 517 VAHRSTIRASDLVTLTLDGMAEKGAHAELAKRGLYSLVMSODIKKADQMSMTYS 576
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 631 IAHRTSTIKTANITVYVNGKIAEGTHDELVDGATRKALVEAQRINQKADALEDD 690
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 577 TERKTN--LPLSHVSKISDF-----IDK--AEESTQSKES-----LPEVSLKTL 620
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 691 AEDLTNADIAKIKTASASDLGDKPTTIDRTGTHKSVSALISKRPPETPKYSIMWTL 750
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 621 K-----LNKPEWFFVVLGLTASVLTNGTVHVFESIIFAKITMGC--NNDKITLKHAEIYS 674
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 751 KFAVSFNRREIPYMLIGLVFSVLAGGQGTQAVLAKALSTLSPEOYSKLRHADFW 810
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 675 MIFVLGVICFVSFMQGLFYGRAGEILTMRLHLAFKMLYQDIAMFDEKENSFGGLTT 734
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

Db 811 LMFFVGLIIOFTOSTNGAAFAVCSERLIRARSTAFRTLLKODIAFFDKEENSTGALTS 870
OY 735 ILADIADIQAGTGRIGVLTQNMATNMGLSVIISITYGHEMFFLLSTAPVLAAYGMLET 794
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 871 FLSTETKHLGSGVLTGLTMTSTTLGAATITAIAGIKLALVCIISVVPVLLAGFYRF 930
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 795 AAMTFANKDKQELKHAGKIATEALENIRTVISLREKAFEQMYEMLQTOHRNTRKAAQ 854
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 931 YMLAQFQSRKATAYEGSANFCEATSSIRTVASLRENRVMTIYHAQDAGRTSLISLV 990
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 855 IIGSCYAFSHAFYFAVAGFRFAYLQIAGRMTPDEGMFYTTAAYGAMAIGKTLVLAP 914
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 991 RSLIYASSQALVFPCVALGFWEYCGTLLGHHEYDIFRFVCPSEILFGAOSAGVFSFAP 1050
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 915 EYSKAKSGAHLFALLEKRPNDRSQBEKKRDTCENLFEFEVGFYFCRDRDVLINGL 974
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1051 DMGKAKMAAEFRRLFDRKPQIDNWSSEBEKLETVEGELEFNVRHFRPTREQPVLRL 1110
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 975 SLSEIRGKTVAFVSGSGCKSTSVOLLQRLYDPVOGOVLFEDVDAKELNVQMLRSQIAIV 1034
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1111 DLTVKPGQTVALLVGPSSGCKSTTIALLEFYDAIAGSILYDQKDISKLNINYSRFLSLV 1170
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 1035 POEPLVFNCSIAENIAYGDNRSVPLDEIKEANANANISHFIEGPEKYNTOVGLKAOL 1094
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1171 SEPTLYGTIKENILGLIVEDDVPBEFLIKACKDANIVDFIMSLPEGFNVTVSGKGM 1230
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 1095 SGGOKORLAIAFALLQKPKILLDEATSLDSEKVVQVHADKARTGTCVLVTHRLSA 1154
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1231 SGGOKORVAIAFALLRDPKILLDEATSLDSEKVVQVADAAARGRTTIAVAHRLST 1290
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 1155 IONADLIYVHNGKIKKEGTHOELLRNDRYFKIVNAOSV 1194
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1291 IOKADVIYVFDGKIVESGTHSELVOKGRYTELVNLSL 1330
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT
8
US-09-328-320-2
: Sequence 2, Application US/09328320
: Patent No. 6228615
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Maard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene at/d of
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/328,320
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/996,545
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11766
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763

```



```

Db      170  EVFVYVSVGVFIYGEHLTKQIRRENYLEAILRQNNAYFDKAGEVTRIRIADFNLIQD  229
Oy      115  GIGDKIALLPQNMSTFSGIANGVLYGKMLTYLTSTESPLIMASAAACSRMAYISLTSKUL  174
Db      230  AISEKVGTLTAFATFVAFIYAVYKMYKALICTSTIYALVWVGGSRSFIVKYSKSI  289
Oy      175  SAYSKAGVAEVLSSIRFVAFRAOEKELQRYTONLKDQAFGIKRTIASVYSAGAYVE  234
Db      290  ESYGAGVYAEVVISIRNATAFGTQDLAKQYETHLAEKMWYKQOVLICMIMGMG  349
Oy      235  FMNGTYGAFWYGTSLILNGEPYTIIGVLAFFSVIHSYICGAAPHFETFAIARGAA  294
Db      350  IMFSYVGLGFWMGSRFVYVCKE--VAVGOVLTYLMSILIGSFSLGNVAPRGQAFNTVAAA  407
Oy      295  PHIFOVIDKKPSIDNFSTAGYKPESECTVEKFNVSFNPSRPSIKILKGLIRIKSGET  354
Db      408  AKIYSTIDRSPLDYSDGKVLDFHEGNIFFRNVKHIYPSREYTVMEDVSLSPAGKT  467
Oy      355  VALVGLNNGSKSTVYQQLQRLYDPDQDFIMVDENDIRALNVHRHYDHIGVSGEPVLEGT  414
Db      468  TALVGPSSGSKSTVGLVERFYLPGVGQVLLDGHDIQTLMLRWLQOISLVSGEPVLEFST  527
Oy      415  TISNNIKTG-----RDVTDDEMERARANAADYDFIMEPNKENTLVGEKGAOMSG  465
Db      528  TIFRIEHLGISTKFEHESKDKIRRELYENAMAMNANHDFTMALPEGYDINVOQGRFLSG  587
Oy      466  GOKORIAIARALVNRPKLILDEATSAIDSESKSAVOALAEKASGRRTIYVAHRLSTR  525
Db      588  GOKORIAIARAIYSDPKIILDEATSAIDTKEGVYQALDKAEGRRTIYVAHRLSTIK  647
Oy      526  SADLVTLKDLAKGMAELMAKRGLYYSVMSODIKKADQMSMYSTERTKNSLP  585
Db      648  TANHNVAVMGKIAQGTDELVDKRGTYKLVLEAQRINE-EKEVALADADMDADDFG  706
Oy      586  LHSVYSIK-----SDFIDKAEESTO-----SKEIS-----LPE-----VSILKIILK  621
Db      707  QGVYRIRITAVSSNSLDVADEKARLEMKRTQKSVSAVLSKVPBEPEKYSLMTLVK  766
Oy      622  ----LNKRPMPFVVLGTASVLNGVHPVSTIIFAKIITMGNNDKT--TLKHDAEYISM  675
Db      767  FIGAENRPELAGMLIGLTFESFLAGGQFQAFLYAKAISTLSLPESMFHLRHDANFMSL  826
Oy      676  IPIVILGICFVSYFMOGLFYGRAGELIMRLRHLAFKALYODIAMFDEKENSSTGLTII  735
Db      827  MEFVVGIAQFISLSINGTAFALICSERLIRRARSOAFRSILRODISFDEENSTGALTSF  886
Oy      736  LAIDIAQOGATGSRIGVLTQATNMGLSVIISFYGEMTEFLISIAVLAVTGMIEITA  795
Db      887  LSTETKRLNSGVSGVLTGIIIMTSTLGAAMITLALIGKALVLCISVVFILACGFLIRY  946
Oy      796  AMTGAFANDKQELKHAGKIATALENIRTIYSLTRREKAEQYMEEMLOTOHRTSKAKOI  855
Db      947  MLAGQOORSKASVEGSASVACEATSAIRTVASLTREQDVWGVYHQDLQOGKRSLSIVLR  1006
Oy      856  IGSCTAFSAFIYFAYACFRGCAVILIQAGRMTPEGMFVTFPALAYAGAMAIGKTLVLAPE  915
Db      1007  SSLTAASSQALVFCVALGFMVGGTLLGHHEYSIRFFFCFSEILFEGASASCTVFSFAD  1066
Oy      916  YSKASGAHAFLEALLEKRNIDRSOEGKPPTCCEGNLEFREESFFYPCRPDVFILRGIS  975
Db      1067  MOKANMAAQAFLKLPDSKRTIDINSDEGEKLESMEBEIEFRVYHRYPTRRPEQVYLRGIN  1126
Oy      976  LSIERGKIVAVFGSSGCGKSTVSQQLQRLYDPVOGOVLFDGVAKELANVQMLRSQAIYV  1035
Db      1127  LSVKPGQYIATVGPSSGCKSTTIALERFYDALAGVFWDGDKITLANSYSRFSLSIVS  1186
Oy      1036  QEPVLFNCIAENIYAGDSRVVPLDEIKEAANAANIHSFIEGLPEKTYQVGLKQAQIS  1095
Db      1187  QEPVLYQGTIKENILGVDKDVSEETLLKVCKDANIYDFVMSLPGFDTVVGSKGMS  1246
Oy      1096  GOKORLAIRARLLOPKIILDEATSAIDNDESEKVVQALDKARGRCLVYTHLSAI  1155

```

```

Db      1247  GOKORVAIRARLRDPKVLILDEATSAIDSESEKVVQALDAARGRRTIAVHRLSTI  1306
Oy      1156  QNADILIVLHNGKIKQGHQHOELRRNDIYFKLVNAGSV  1194
Db      1307  QNADIIYFDDGKIVESGTHHELIRKGRYVELVNLQSL  1345

RESULT 10
US-08-395-246C-2
; Sequence 2, Application US/08395246C
; Patent No. 5773214
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,246C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35784
; REFERENCE/DOCKET NUMBER: x9683
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; TELEFAX: 317-277-1917
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1307 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-395-246C-2

Query Match 36.0%; Score 2166; DB 1; Length 1307;
Best Local Similarity 38.3%; Pred. No. 5,2e-187;
Matches 475; Conservative 243; Mismatches 464; Indels 58; Gaps 12;

Oy      6  LASTVAGACLPLMPILVIGMSDNLISGLVQNTYSF-----RLTLVYVIGVALVIFG  60
Db      67  LAIATAGGAALPFTYLGKFT-STFODIATHRITTYHFMHETLKNVYFIYGAABFVAL  125
Oy      61  YIOISLMTITTAQRTRIKRQPFHSHVLAODIGMFDSCDIGELNTRMT-DIDKISDGIK  119
Db      126  YLATVGFITYTGDHVQIQINVEYFOAILRQNIAPFTLGAEGEITRTLTADTNLIQDISK  185
Oy      120  IALLPQNMSTFSGIANGVLYGKMLTYLTSTESPLIMASAAACSRMAYISLTSKUL  179
Db      186  VGLATLGLSTFVYAFITAIKMKKALLISASLALLTMGGCSTLMLFSKKALEYQGR  245
Oy      180  AGAVAEVVISIRFVAFRAOEKELQRYTONLKDQAFGIKRTIASVYSAGAYVEFNMG  239
Db      246  GASMAEDIIDSTRTVAAPNAGETLARKYESHLKADAGPCKSKVITRAINWALCLCTMYLN  305
Oy      240  YGLAFYGTSLILNGEPYTIIGVLAFFSVIHSYICGAAPHFETFAIARGAAPHIQ  299
Db      306  YGLGFWMGSRFVYVCKE--VAVGOVLTYLMSILIGSFSLGNVAPRGQAFNTVAAA  407

```



```

QY 300 VIDKKPSIDNESTAGYKPESEIEGVEFNVSFNPPSPRSIKLKLNDLRKSGEFTVALVG 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 366 TIDROSPDALSDGOKTLEFGNINIVLQNIHRVYPSREPVTAHDLSCYIPAGKTTAFVG 425
QY 360 LNSGSKSVVOLLQRLVYPPDOGFIVNDENDIRALNVRRYRHDIGVSEPVYFGTTISNN 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 426 PGSGSKSTISLSEFYDPVAGTITMLDGHDIOTLNLRLMROOMSLVSGEPRLFATTIAEN 485
QY 420 IKYG-----RDDVYDE---EMERARANANADFIEMEPNKNFTLVGKGAOMSGGOKOR 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 486 IRYGISRFEKSESTYELRKNVEAARANAADFIEMALPNCYDTNI--ESFSLSGGOKOR 543
QY 471 IAIARALVNPRIKILIDRATSALDSEKSAVOALEKASKGRTTIVAHRLSTIRSAOLI 530
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 544 IAIARAIKDKPRILLDDEATSDLTOKSEKLYOALDKASKGRTTIVIAHRLSTIOKAYNI 603
QY 531 VTIKDGMAEKGAAHELMAKKGLYSLVMSODIKKADQEMSMYSTERKTNLSPLHS-- 588
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 604 IYLANGQIVEGGPHEHMDRGRICYDMVEAHEIKRYSRY-SKRRYS-OLTLNLSFKHNP 661
QY 589 -----VKSIKSDFIKAEESTQKESLSPEVSLKILK----- 621
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 662 TFFPKDVPGDDESIDYIILSDASDGLHTEGKORPVSRMSLSHLMQPVKEANSFWTL 721
QY 622 -----LNKPEWPFVYIGTSLASVINGTVHPVFSIIFAKIIIMFG--NNDKTLKHAELY 673
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 722 FKFLASFNRPEWPFLLGLGASIIAGIOPSGAVLFAKAVSTLSLPLEYPRLRDANFW 781
QY 674 SMIFVILGVICFVSIFMGLPFRAGGELLTMRRLHAKRAMLYODIANFDEKENSIGIT 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 782 CLMFMIIGIVSLVYSGVTLFAYSEKMYRARSQAERVIHODISFPDQDENTGALT 841
QY 734 TILAIDIOIGATGSRIGLTONATNMGLSVIISIFIGEMETFFILSIAPLAATGME 793
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 842 AFLSGTEHLEIGISVITGLTIIYVSNLVASIGVALVIGMKIALVCISAVPMLMCGFYR 901
QY 794 TAAMTGAFANKDQKELKAKGIALEALENIRITVSLTRREKAEQOMYEMLOTOHRTSKA 853
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 902 VWMLERFORAKKAYQESASSACEAASAIRTVVSLTMEETALOSYQAOURLQKSDILPI 961
QY 854 QIIGSCYAFNAFIYFAAAGFRCAVYLQAGRMPEGMFIYFALVANGAIGTIVLA 913
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 962 VKSSLVYASSQALPFECALGFWYSGSLHGEYSLFOFYCFSEVIGCAAGTVFSHA 1021
QY 914 PEYSKAGKAAGHLFALKEKKPNIDRSQEGKKPDCGEGLEREVSFYPCPDVITLIG 973
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1022 PDMGAKAH-AAREFKRLSSDTMAHSRSKGVYVTSMRGLVEFRDVSFRYPSLEQPILRH 1080
QY 974 LSLIERGKTYAVGSSGCGKSTVOLLQRLVDPVOGVLEFDVAKELNVQMLRSQIAI 1033
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1081 LNLTKPGQFVALVGASSGSKSTTIALLERFYDPLKGGVYDGNKIIITLEMSTRSHLAL 1140
QY 1034 VPQEVLEFNCISAEINAVGDSRVPLDEIKENANANIHSTIEBLPEKYNTQVGLKAO 1093
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1141 ISOEPTLFGOTIRENILLGSMTPHATDPLVACAKDANIYFELISLPGCFNTIVGNKGM 1200
QY 1094 LSGGOKRLATARALLOKPKITILDEATSAIDNSEKVOAHLDKARFGKTCVLTVHRLS 1153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1201 LSGGOKRIATARALIRNPKITILDEATSAIDSESEKVOAALDAAANGRTTIVAHRLS 1260
QY 1154 AIQNAIDLIVLHNGKIKBOGTHOELLRNRRDIYFKLVNOS 1193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1261 TIQRADLIYLDQGEVSGTHRELLRKKGRYELVHLQ 1300

```

```

? TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
? TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
? NUMBER OF SEQUENCES: 3
? CORRESPONDENCE ADDRESS:
? ADDRESS: Patent Division/AEH
? STREET: Lilly Corporate Center
? CITY: Indianapolis
? STATE: Indiana
? COUNTRY: USA
? ZIP: 46285
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/612,521
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Hamilton, Amy E
? REGISTRATION NUMBER: 33,894
? REFERENCE/DOCKET NUMBER: X-9693
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 317-276-1294
? TELEFAX: 317-276-3169
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1408 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-612-521-2

Query Match 35.2%, Score 2116, DB 1: Length 1408;
Best Local Similarity 38.3%, Pred. No. 2,1e-182;
Matches 482; Conservative 241; Mismatches 461; Indels 76; Gaps 20;

QY 1 MITGLIASLVNACCLPMPVIVGEMSDNLISGCLVQNTYSPFRLT----- 46
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 156 MYGLIIVLVNAGSCQPLMTLLIFGRLLTSFTNAYI-ANQISOGGLTPETSAALQAKDL 214
QY 47 -----LYVVGIVVAALIEGYIQISLMIITPAARQTRIRKOFHVSVAODIGWFDSCD 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 KTQSGHNLVYMAIGIGFLATWLMFTMNTYTGELNSRIRERYLAALVROEIAFDLGG 274
QY 99 IGEINTRM-TDIDKISDGIKIALLFQNMSTFSGIAGVGLVKKKLTLYLTSPLTMA 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 275 AGEVATRIOTDCHLVQEGTSEKVALVFQAGTFVCGVLAERSPRLAGALVSLIPYIML 334
QY 158 SAAACSRVYISLSEKELNYSKAGVAEEVLSITVYAFRAQEKEL-QRTONIKDKDKD 216
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 335 CGGIMTMAKKYGTALDLHIAKAGSLAEVIGSITVQAF-GKEKIIADKADHDIEQSKI 393
QY 217 FGKIRTIASKYSLSGAVYEFNMGTGIAFWYGSLLINDEPGYTGTVLAVFFSVYHSYC 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 394 VGRKSIIEDEGLSIMEFVITYAVALAFYGGIIVSNQA--DSGIVINVMSTILISGFS 451
QY 277 IGAAPVHEFTFAIARGAAPHIFQVTDKRPISIDNESTAGYKPESEIEGVEFNVSFNPSR 336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 452 NMLAPPELAAYTKARGAAGAKLFATIDRVPAIDSAESEFKPDGLGELISFEVYKFPYPSR 511
QY 337 PSIKILKGLNLRKSGEVALVGLNGSKSVYVOLLQRLVPPDOGFIVNDENDIRALNV 396
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 512 PSIPILKGFYTFTEAGKFFALVAGSGSKSTVSLIERFYDPSVNVLDGIRSLNLN 571
QY 397 HYRDIIGVSOEPLYFGTTISNNIKYG-----RDDVYDEE---EMERARANANADFIEM 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 572 WLROQIGLYSGEPTLFGFTYKGNVHGLIGSRVKNASLEKFEYLVKACVDANHNFMK 631
QY 448 FPNKNFTLVGKGAOMSGGOKRIATARALVNPRIKILIDRATSALDSEKSAVOALEK 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
US-08-612-521-2
: Sequence 2, Application US/08612521
: Patent No. 5786463
: GENERAL INFORMATION:
: APPLICANT: Peery, Robert B
: APPLICANT: Skatrud, Paul L
: APPLICANT: ThorneWell, Susan J

```

```

Db 632 LPQGYDVMWGERGMLSGGOKORVAIAIAVSDPILLLDEATSDALDQSEGIYODALDK 691
OY 508 ASKGTITVAHRLSTIRSAOLIVLKGOMLAEKGAHAEMLA-KRGLYYSTYMSODIK-- 564
Db 632 ASRGHTTITIAHRLSTIRSDRIYVGGEEVLEQSHNDLANENGPAQVNNOKLAGE 751
OY 565 -----KADEOMES-----MTYSTERKTNSLPLH-----SVKSIKSPFID--KAEEST 604
Db 722 AAALALQVDDDDLEDDDAVFIQSSPMQEKQOLHRAVYGRSLASTAMDIAKRAEEVA 811
OY 605 OSKEISLPEVSLKLIKLIKPEMPFVVLGTASLVNGVHPVFSIIFAKITMFGNDKT 664
Db 812 GEDKIPSSGGLYARLLRMSADKFIYIIAFAIACAGMYPSLALIFGKALSDFEIQDPA 871
OY 665 TLKHAEIYSMLFVLGVICFVSYFMQGLFYGRAGEILTMRLHIAFKAMLYODIAMPE 724
Db 872 ELRHALLSRSAWYFTTALAALAAAFVIFQSGFSRAGMDLVGRKLFYTLHHDIEWPE 931
OY 725 KENSTGGTLTTLAIDIAOIGATGSRIGVLTQNTMGLSVLITISFYGEMTFILISAP 784
Db 932 EKNSTGATVSNLADQPOKVOGLFGLTIGVYVQSCATLIGGCIIGLCYGLLALIGIACIP 991
OY 785 VLAVTGMETAAMTGPAKNDQELK-HAGK--IATEALENITIVSLTREPKEQWYEM 841
Db 992 ILVSGYIRLKV--VLKQDMKKLHAASAHLASAAGAVKTVASLTFREKDVRIYSEA 1048
OY 842 LOTQHR---NFSKAOIIGSCAFSAHFIFYFAAGFRGATLLOAGRTPEGMFTYFPA 898
Db 1049 LKAPKMLNFRISIKSOCL--FAASOGLETCIIALVFYIGALMIIDAKSTASFTVLNS 1105
OY 899 IAYGAMAIKTLVLAPEXSKAKSAGAAHLFALLEKPNIDRSQOEKPPD--TCEGNLEFR 956
Db 1106 IYFASIQGNVTFYFPDASKANSASAFRSIDNEPAIAEENEGEVLDDHKVYGVRIE 1165
OY 937 EYSFYPCHRPDVFILRGLSLSTERKTYAVFVSSGGCKSTVOLLQRYLDPVOGYLFDG 1016
Db 1166 GVHFYFPRPGVRLNLTIDVPAGTYVALVGPSCGSKSTTIOMLERFYDLAAGRVLTDG 1225
OY 1017 VNAKLNQMLRSQAIYVQDEPVFNCSIAENIAYGDNRSV--VPLDEKEAANAHTS 1074
Db 1226 IDIKELNLAYSQSISLVSQEPFLVAGTIRFNILLGANKPIEYVODEIDAACKDANIYD 1285
OY 1075 FIEGPEKNTOVGLKGAQSLSGOKORLAIARALLQPKILLDEATSDALDSEKVVYOH 1134
Db 1286 FIVSLPDEFTYVGGKQSGOKORLAIARALLRNPVLLDEATSDALDSEKVVYOE 1345
OY 1135 ALDKARTGTCLVYTHRLSAIONADLIVLHNGKIKEQGTQHOELLRNBDIYFKLVNAQSV 1194
Db 1346 ALDKAKAGRTTIAIAHRLSSISQHSRITYFSEGRVAEHGTQHOELLAKKGYYELVQOMNL 1405

```

```

RESULT 12
US-08-996-644-2
: Sequence 2, Application US/08996644
: Patent No. 5945324
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Waard, Maarten
: TITLE OF INVENTION: Multiple Drug Resistance Gene atloc of
: TITLE OF INVENTION: Aspergillus nidulans
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESS: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/996,644
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11765
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1308 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-996-644-2

```

```

Query Match 28.2%; Score 1695.5; DB 2: Length 1308;
Best Local Similarity 33.9%; Pred. No. 2,5e-144;
Matches 403; Conservative 243; Mismatches 479; Indels 65; Gaps 19;

```

```

OY 47 LTVYVIGVALLFGYIQLSLMTITTAARQTKRIKOPFHSVLQDIDQMPSCDIGELNTRM 106
Db 141 LFFVILGIRLVLTSTYNTLLTYAATRYRNRHAYLRKALISQEVAYYFPGSGGSTIAQA 200
OY 107 TDIDK-ISDGIQDKIALLFQNNSTFSIGLAVGLVKGKLLTV-----TLSTPLMAS 158
Db 201 TSNGLIQAGADKIGLIFQGLAAFTVLSLSRLCMKWLTLICICIPATVIGTGVAAV 260
OY 159 AAACSMYISLSKELSAKAGAAVEVLSIRYIARAOEKELQRTQNLKAKDNG 218
Db 261 EAGHETRLQI-----HAQANSFAGILAGKVAHAGMDSLVKRPDEYLVEAHKV 313
OY 219 IRTIASKVSLAAVYFPMNGTGLAWYGTSLILNGEPYTGTVLAVFYSYHSYCIG 278
Db 314 KTIPLGLFSAEYIITLYGLAFWQIHMGKEIG-TAGDITVLLSVYIASINTL 372
OY 279 AAVPEFPAIRGAFAHIFQVYIDKKPSIDNFTAGYKPEISIEGVFEKNSFNPSRPS 338
Db 373 ILAPYSIERSRAASAAQFLRIDRESEINPGKEBLEPRLVDVLELVNTPFSYTRRG 432
OY 339 ITILGLNLRITSGETVALVGLNGSGKSTVOLLQRYLDPDGFIVDENDIRALNRYH 398
Db 433 IYVLDNFSLSKVPAGKYATLVAGSGSGKSTYVGLLEKWNYPSTGALRDLNLSLWGMV 492
OY 399 RDHIGVSGEPVLFQGTITINNHIKYG-----RDDYVDEEMER--AAREANAVDFTEBP 449
Db 493 RNRVRLVQDEPVLFGQSVVDNIRYGLVGPWENASREBOMERYQEAKLAVHFEISETL 552
OY 450 NKEFTLVGEKGAQMSGQKORLAIARALRNPKILLDEATSDALDSEKSAVOALEKAS 509
Db 553 DGYDITLIGRGGLISGGQORVAIAASVVSQPKVLLDEATSDALDHAFTIYQKALDKAA 612
OY 510 KGRITIVAHRLSTIRSAOLIVLTKDGMIAEKGAHAEMLAKRGLYISYMSODI----- 563
Db 613 EERTTIVIAHRLATIRKANIVYMSKGHIVEQGTESLAKGVYVAGLVKQIONLAVNAGA 672
OY 564 -----KADEOMESMTYSTERKTNSLPLHSVKSISKDFIDKAEESQSGEISLPEVSLK 618
Db 673 HDNVMEGEGEDVALLETETAVTRPT-SIKRANSTDR-DDEVNKHMDM--LAALA 728
OY 619 ILKLNKPEMPFVVLGTASLVNG-TVHPVFSIIFAKITMFGNDKTTLKHDA-----BI 672
Db 729 YLVRECPELKMAVLYLVLLCIGGCANYPGQATILMSVVEVF-----TLSGAMLDKGF 782
OY 673 YSMIFYILGVCIFYFMQGLFYGRAGEILTMRLHIAFKAMLYODIANFDEKENSTGGL 732
Db 783 YASMLVLAAGLICVLAAGYATNTIAQHLSHWFRLILHMDLRDIOFFDEENNTGAL 842
OY 733 TITLAI-----IAIQGATGSRIGVLTQNTMGLSVLITISFYGEMTFIL-STAPL 786

```

```

Db      843  --VSRIDSYPHALIELMAGYNIALVYIAVLQVYTCG---ILAIASFSGKGLVVFEGGIPPL 897
Oy      787  AVTGMIEETAAMTGFANKDKOLKHAGKLTATEALENIRTVISLTREKAEOMYEEMLOTOH 846
Db      898  VGAGVNRIRVDSRLDRQTSKRTGTSSTIASAEVNAIRTVSSLAIEYVLRRTTELDHAV 957
Oy      847  RNTSKKAOIISCYAFSHAFTYFAVAAAGFRGAYLIQAGKMTPEGMFTVFTALAGAMAI 906
Db      958  SSSVAPMAATMFCGLTQCIETWFOALGEFWGCRLVSLGETSMYSFFAFLSVFPAQAS 1017
Oy      907  GKTIVLAPEYSKAKSGAHLFALLEKKNIDRSOEGKKPDTCECNLEFREVSFFPCRP 966
Db      1018  AOLFWMSITIKGINATNATYIAMLHOLQPTVRETPENHDKPGSGAPIMADVNRFSYPIRP 1077
Oy      967  DVFIIRGLSLSTERGKTVAFVSSGCGKSTSVOLLQRLYPDQOVGLDPDGVAKELNVQW 1026
Db      1078  DAPILKGVNLIKNGQFIATFVSSCGCKSTMAMLEFRYDPTGSTITDASTLDINPIS 1137
Oy      1027  LRSQIAIYPOEPLVFNCSIAENIAYGDSRVVPLDEIKEMANANIHSTIEGPEKYNTO 1086
Db      1138  YRNIAVALVOEPTLFQGTIRDNISLSDAVKSVSDEQIESALRAANAMPVSSLPGIYTP 1197
Oy      1087  VGLKGAQLSGGOKORLAIRALLOKPKILLDEATSALDNSEKVVQHALD-KARTG-RT 1144
Db      1198  AGSGGSQLSGGOKORLAIRALIRDPKILLDEATSALDTESEKIVQKALBGAARDGDL 1257
Oy      1145  CLVYTHRLSATQNDLIYVNLHNGKIKEOGTHOELLRNNDIYFKLVNASV 1194
Db      1258  TVAAVNRHSLTKIDANVICFFGCKIAEMGTHOELLIRGGLYRRMCEAOL 1307

```

## RESULT 13

US-09-352-552-2  
Sequence 2, Application US/09352552

Patent No. 6060264

GENERAL INFORMATION:

APPLICANT: Skatrud, Paul L.

APPLICANT: de Waard, Maarten

TITLE OF INVENTION: Multiple Drug Resistance Gene atcc of

TITLE OF INVENTION: Aspergillus nidulans

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/352.552

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/996,644

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-11765

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

TELEFAX: 317-276-2763

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1308 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-352-552-2

Query Match 28.2%, Score 1695.5; DB 3; Length 1308;  
Best Local Similarity 33.9%, Pred. No. 2,5e-144;  
Matches 403; Conservative 243; Mismatches 479; Indels 65; Gaps 19;

```

Oy      47  LYYVIGVAAALIFGYIQSLMTWITTAAROTKRIKRFQFHSVLAODIGMDSODIGELNTRM 106
Db      141  LYFVIGIARLVLTSTYTWLTLTYAAYRIVNRIRHAYLKAALSQEAVAYDPFGSGSIAQA 200
Oy      107  TDIDK-ISDGIGDKTALLFQNNSTFISGLAVGKGMKFLV-----TLSTPLMAS 158
Db      201  TSNKGLIOAGASDKIGLFGGLAAAEVTLTSLRMCKWKLTLICICIPATVITGTTGVAAY 260
Oy      159  AAGCRMYISLTKSLASVSKAGNAAEVLSIRIVIAFRQEKELQVTONLQAKPFG 218
Db      261  EAGHETRILOI-----HAQANSFEGILAGVKVHAFGMKDSLVRFDELYEAHRYG 313
Oy      219  IKRTIASKVSIGAVYFMNGTGLAFWYGTSLILNGEGYTTIGTVLAVFFSYHSSYICG 278
Db      314  KKISPLGLLSAEYTTIYLGGLAFMGIHMFGEIG-TAGDLFTVLLSVIASINLT 372
Oy      279  AAVPHEFTFALARGAAPHFOVIDKRPKSIDNESTAGYKPESTIEGTEKKNVSFNPSRPS 338
Db      373  LLAPISIEFSRASAAQDLFRIDRESEINPYKGEPLERVLGDVELENTVFSYTPRPG 432
Oy      339  IKILKGLMDRIKSGFTVALVGLNGSGKSTVVOQLRLYPDQDGFIMVDENDIRALNVKH 398
Db      433  ITVLNFSILKVPAGKVTALVQSGSGKSTIVGLLERWNPSTGALRLDGNLISELNVWL 492
Oy      399  RDHIGVSOEPLFGTSTNNIKYV-----RDVTDSEMER--AREANAYDETFMEFP 449
Db      493  RNRVNLVOOEPLVFGSVFDNIRYGLVGTPMENASREOMERVQAAALAAHETSILT 552
Oy      450  NKFNLTVEKGAGMSGOKORAIARALVRNPKILLDEATSALDSEKSAVOALAKRAS 509
Db      553  DQYDILIGERGLLSGGOKORAVIARSVQPKVLLDEATSALDPHAEFTIVQKLRKA 612
Oy      510  KGRITIVVAHRLSTRSADLYTLKDGMLAEKGAELMAKRGLYSYLSMSODI----- 563
Db      613  EGRITIVIAHKLATIRKADNIVMSKGHIVEGTHESLIADGVYAGVAKVLQNLVNAS 672
Oy      564  -----KKADEQMSMTYSTERKTNLSPLHSYSIKSDEIDKAEESTQSEISLPEVSLK 618
Db      673  HDNVNEBGEVVALLETETRAVTRYPT-SIRGRNNSIKDR-DOVENKHKHDM--LAALA 728
Oy      619  ILKLNKPEMPFVYVIGTLASVLNG-TVHVPFSITIEAKITTFMGNDKTKLHDA-----EI 672
Db      729  YLVRECPELKNAVYLVVLLGCLGCGCAMYPGQALIMSRVEVF-----TLSDGMDLKGDF 782
Oy      673  YSMITVILIGVICSFYFPMQGLFYGAGELIMRLHIAFLKAMLYODDIAMPDEKESUGL 732
Db      783  YASMLIVLAAGCLICYLVAGATNTIAOHLSHWFERLLIHLHDLRODIOFPFREENTTAL 842
Oy      733  TTIILID-----IAOIGATSGRIGVLQONATNMGLSYISFIYEMETFTLL-SIAVPL 786
Db      843  --VSRIDSYPHALIELMAGYNIALVYIAVLQVYTCG---ILAIASFSGKGLVVFEGGIPPL 897
Oy      787  AVTGMIEETAAMTGFANKDKOLKHAGKLTATEALENIRTVISLTREKAEOMYEEMLOTOH 846
Db      898  VGAGVNRIRVDSRLDRQTSKRTGTSSTIASAEVNAIRTVSSLAIEYVLRRTTELDHAV 957
Oy      847  RNTSKKAOIISCYAFSHAFTYFAVAAAGFRGAYLIQAGKMTPEGMFTVFTALAGAMAI 906
Db      958  SSSVAPMAATMFCGLTQCIETWFOALGEFWGCRLVSLGETSMYSFFAFLSVFPAQAS 1017
Oy      907  GKTIVLAPEYSKAKSGAHLFALLEKKNIDRSOEGKKPDTCECNLEFREVSFFPCRP 966
Db      1018  AOLFWMSITIKGINATNATYIAMLHOLQPTVRETPENHDKPGSGAPIMADVNRFSYPIRP 1077
Oy      967  DVFIIRGLSLSTERGKTVAFVSSGCGKSTSVOLLQRLYPDQOVGLDPDGVAKELNVQW 1026

```

Db 1078 DAPIKGVNLKINKQPIAFVSSCGCKSTMTAMLERFYDPTGSLTIDASTLTIDINPIS 1137  
Qy 1027 LRSQIAIVPOEVLVFNCSIAENIATGDSRVVPLDEIKEANANINHSIEGLPEKYNTQ 1086  
Db 1138 YRNIVAVQOEFTLEFGOTIRDNISIGDAVKSVSDEQISALAAANAMPVSSSLPOGIYTP 1197  
Qy 1087 VGLKGAOLSGGOKORLAIARALLOPKLILLDEATSAIDNSEKVVQAHLD-KARTG-RT 1144  
Db 1198 AGSGSQSLSGGOKORLAIARALIRDPKILLDEATSAIDTESEKIVQKALGEGARDGRL 1257  
Qy 1145 CLVYVHRLSAIONADLIYVHNGKIKEDGTHQELLRNNDIYFKLVNAOSV 1194  
Db 1258 TVAVAHRLSTIKDANVICVFEGGKIAEMGTHELLVIRGSLYRRMCEAQL 1307  
RESULT 14  
US-08-232-537-2  
: Sequence 2, Application US/08232537  
: Patent No. 5516655  
: GENERAL INFORMATION:  
: APPLICANT: Peery, Robert B.  
: APPLICANT: Skatrud, Paul L.  
: TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
: TITLE OF INVENTION: AUROBASIDIUM PULULANS  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Eli Lilly and Company  
: STREET: Lilly Corporate Center  
: CITY: Indianapolis  
: STATE: Indiana  
: COUNTRY: USA  
: ZIP: 46285  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/232.537  
: FILING DATE:  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Plant, Thomas G.  
: REGISTRATION NUMBER: 35,784  
: REFERENCE/DOCKET NUMBER: X9212  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 317-276-2459  
: TELEFAX: 317-276-1917  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1302 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: US-08-232-537-2  
Query Match 27.7%; Score 1664; DB 1; Length 1302;  
Best Local Similarity 32.1%; Pred. No. 1.8e-141;  
Matches 400; Conservative 258; Mismatches 503; Indels 84; Gaps 22;  
Qy 4 GILASLVNACGLPLMLVLGEMSDNLISGLVQNTYSPFRLLTYVVGIGVALLFGYIQ 63  
Db 87 GILAIAMGCPILIGIYFGQITISF-----PPEDVLRLRYOL-VGVAC---GYFI 135  
Qy 64 IS-----LWITAAQTRKIRKQFHSVLAODIGMFDSCDIBELNRMDIDIKISGID 118  
Db 136 VTTGTAIAMGLTGEKISRFRFETLVERLIGLEQAFDIDPITNLTEKIEADIQITSE 195  
Qy 119 KIALFQNMSPFSIGLAVGLVKGKLLTVTLSTSPILMASAACSHPVLSISKELAYS 178  
Db 196 KVGITIGSISYVAAPFYIGFILNAKLT-----GILFAAVIPIMALLIYTVSSRIAKTT 248

Qy 179 K-----AGAAVEEVLSSIRPTVIAFRAQEKELQRYTONLMDAKDPGIKRTIASVSG 230  
Db 249 KATEYTEAAGRIKAEIAHAAVYVQAFGAENLSKHEHYLLKISARYALRKSSAFLMIG 308  
Qy 231 AAYFFNGTGYGLAFWYGTSLINGEPG-YTIGTVLAVFFSVIHSYSCIGAAVPHFTFAI 289  
Db 309 LYFFTAYSANALAFWEGSLA--AESGNNAGVYAVVLELIDASTVVOQFQFGLSFT 366  
Qy 290 ARGAAPHIPQVIDKPS-IDNESTAGY--PESIEGVEFKVSNFVPSRPSTIKILGIN 346  
Db 367 AAAAGSEYVEILNHPOSEINIVSSEAGQEAATESMKADLVFRNTEYVYPARTSARALEMS 426  
Qy 347 LRIKSEETALVGLNNGSGSTVVOQLRLYDPDPDGTIMDENDIRALNRHYRHDHIGVS 406  
Db 427 LILKAGQMAIVGTSGCGKSTVSLRLRYLIDSSGQLTGSHDKIDPNRSLKRYTALVD 486  
Qy 407 QEPVLEGTITISNNIKY-----RDDVTDEMERAREANAYDIEFPNKFNTLVGEGK 460  
Db 487 QDSVLFSSSVLENISYGLDEHSLDDVLERCLEAKKAN-IDFVDFLPOGITHRTGNG 545  
Qy 461 -AOMSGOKORLAIARALVNPRLILLDEATSALDSEKSAVOALEK-ASKGRTIYVA 518  
Db 546 YTSLSGGQNRICLALRALVKKPALLLDEPTAALDANSEGLIMDAVSAATGTIVVVA 605  
Qy 519 HRLSTIRADLYTLKDGMLAEKGAHAELMAKRGLYYSIVMSODIKKADQMSMTYTE 578  
Db 606 HRLSTVSDSPNIVLMGAVIEGNDHELMQEGAFNLIOQOLNDADDESSAEVAAT 665  
Qy 579 RRTNSLPLHSVKSIRKDFIDKAEES-----TOSKEISLPEVS-----LKLKLINKPE 626  
Db 666 SQYT-----PQKASKSE--DSAASDTEYVPPAKKEDKPAKKAGFWKLLRLRLAKSD 718  
Qy 627 WPFVVLGTLASVNGTVHVFHSIIFAKIITMGNDKTKLKHDAEIVSMIVLIVYICV 686  
Db 719 SPILAGLAASIYSGGILIGEAIVFNLSLVNLDSPFRSRADJFLFLFILLALIF 778  
Qy 687 SYFMQGLFYGRAGEIITMLRLHAFKAMLYODIAMDDEKNSGGITLAIADIAIOGA 746  
Db 779 SYAGNCCFGIVSSHFAVIOHISLASILRQDMQWSSG--SVPSLMSLSISDAGGLACT 836  
Qy 747 TGSRLGVLQNMNTMGLSVIISPIYGMETPLISTAPVLAIVGMETAAFTAFKDKO 806  
Db 837 SGVAIGTIFTVCVSTIGGILLAVVAMKIAVVLAAVPMITTAGYVRLRLALAESRHS 896  
Qy 807 ELKHAGKIALENIIRITVSLTREKAFPEOMYEMLQTOHRNTSKKAOIIGSCYAFSHA 866  
Db 897 AVNDASIAAECRGIRTIASIGREGVSRASNAVKEPYDKIRFTLITNTLLALSFSI 956  
Qy 867 IYFAVYAGFRFAGYLIQAGRMPEGMFYFTALATAGMAIGKTLVLAPEYSKAKSGAHL 926  
Db 957 TYFVALAYMWGAKOVNRNGTYSOLDFFIVLPALEFSAQSAGQJFSLSPESRAGVARNV 1016  
Qy 927 FALLEKPP--NIDSRQ-----EGKKRDTCGNELEFRVSPFYRCRDPVFI 970  
Db 1017 FGLHDQKPTIYVDNAKOSALBSSTISITPLEKASPSGGAWEFKNVSICLYSKQHRA 1076  
Qy 971 LRLGLSIEKGTAVFVSSGCGKSTSVOLLRLYDPVOGQVLFVDVDAKELNVOMLRSD 1030  
Db 1077 LQNVNISIRGEPIALVPGSAGKSTILLSLQRFYDPTGASVOLDGODIREVAVPHQRGR 1136  
Qy 1031 IAIVPOEVLVFNCSIAENIATG-DNSRVVPLDEIKEANANINHSIEGLPEKYNTQVL 1089  
Db 1137 LGLVPOEPLFPGISISYVIGLGAAPQULYTRDIEIKACGIIHEPIMSIPREYSTECGT 1196  
Qy 1090 KGAOLSGGOKORLAIARALLOPKLILLDEATSAIDNSEKVVQAHLDKARTRCTCYVT 1149  
Db 1197 NSKLSGGOKORLAIARALIRSPVLLDEYTSALDAHSEQOIKKAVDCAASVDRITIVVA 1256  
Qy 1150 HRLSAIONADLIYVHNGKIKEDGTHQELLRNNDIYFKLVNAOSV 1194  
Db 1257 HRLSTVQNMADRIYVFPDGRVVEGSHAEVLVAGGLVAGVLAQTLL 1301



---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:32:43 ; Search time 61.37 Seconds

(Without alignments)  
1442.359 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012  
Sequence: 1 MIIIGILASIVNGACLEPLMLP.....QELLRRNDYFKLVNAQSVQ 1195

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_1101:.\*  
1: /SID58/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SID58/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SID58/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SID58/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SID58/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SID58/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SID58/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SID58/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SID58/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SID58/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SID58/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SID58/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SID58/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SID58/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SID58/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SID58/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SID58/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SID58/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SID58/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SID58/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SID58/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SID58/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3557	59.2	1281 22	AAE00310
2	3556.5	59.2	1283 22	AAAB81065
3	3554	59.1	1280 18	AAW44073
4	3554	59.1	1280 21	AAAY58186
5	3554	59.1	1280 22	AAAB81959
6	3554	59.1	1281 22	AAE00308
7	3554	59.1	1281 22	AAE00309
8	3549	59.0	1280 22	AAAB81068
9	3549	59.0	1280 22	AAE00304
10	3549	59.0	1281 22	AAE00303
11	3548	59.0	1280 22	AAAB81064

12	3545	59.0	1280 21	AAV58187
13	3545	59.0	1280 22	AAAB81066
14	3545	59.0	1280 22	AAE00306
15	3543	58.9	1280 8	AAV70452
16	3540.5	58.9	1272 21	AAV70597
17	3539	58.9	1280 11	AAAR04868
18	3538.5	58.9	1272 21	AAV70596
19	3537.5	58.8	1279 22	AAAB81067
20	3537.5	58.8	1279 22	AAE00307
21	3535	58.8	1280 14	AAAR4297
22	3534	58.8	1280 15	AAAR63624
23	3531	58.7	1280 22	AAAU04347
24	3528.5	58.7	1279 19	AAW80294
25	3527	58.7	1276 14	AAW48997
26	3526.5	58.7	1276 14	AAAR35199
27	3526.5	58.7	1276 21	AAV58188
28	3522.5	58.6	1276 21	AAV58189
29	3521	58.6	1280 19	AAW48998
30	3521	58.6	1280 19	AAW48999
31	3517.5	58.5	1275 22	AAAB60409
32	3517.5	58.5	1275 22	AAW49000
33	3515	58.5	1280 19	AAW49000
34	2367.5	39.4	1275 20	AAW82594
35	2282	38.0	1334 20	AAV02630
36	2273.5	37.8	1349 20	AAV16434
37	2181	36.3	1333 22	AAAB62495
38	2166	36.0	1307 17	AAAB92255
39	2153	35.8	1243 21	AAAG38102
40	2153	35.8	1286 21	AAAG38101
41	2134.5	35.5	1205 21	AAAG39103
42	2130	35.4	1254 21	AAV92173
43	2116	35.2	1408 19	AAAB62871
44	1916	31.9	1261 22	AAAB62493
45	1869.5	31.1	1402 22	AAE02402

#### ALIGNMENTS

RESULT 1	
AAE00310	
ID	AAE00310 standard; Protein: 1281 AA.
XX	
AC	AAE00310;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Dog P-glycoprotein (PGP) allelic variant (Genotype D).
XX	
KM	Dog: P-glycoprotein allelic variant; PGP: multidrug transporter;
KM	MDR1; drug bioavailability; transgenic animal; genetic model.
XX	
OS	Canis familiaris.
XX	
Key	Location/Qualifiers
FT	Misc-difference 25
FT	/note= "Asn of Genotypec substituted by Lys"
FT	Misc-difference 197
FT	/note= "His of Genotypec substituted by Gln"
FT	Misc-difference 329
FT	/note= "Ser of Genotypec substituted by Thr"
FT	Misc-difference 1148
FT	/note= "Met of Genotypec substituted by Val"
XX	
PN	WO200123540-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26767.
XX	
PR	28-SEP-1999; 99US-0156510.
XX	
PA	(GENT-) GENTEST CORP.

[illegible]

QY	519	HLRLTINSADILVTLLKGMALAEKAAHELMARAGLLYSLYMSDDIKKADQEMESMTYSYE	578
Db	589	hrlslvtrnadiaagfdggvivekqnhkbelmkckgylyfklvtlmg-trgnatelenagcsk	647
QY	579	RKTNSLPRHSAVSKISKSPFI-----DKAEESTOSKEISLEPVSLLTKILKAN	623
Db	648	sesadalen-spkdsgssalikttrsttrrsihppgqgdrtlgkdelnlenyppvstwlrlkln	706
QY	624	KPEMPEVVLIGTLASVLAVGVHPHFSITIAKIITMF-GNNDKTTLLKHDAEITYSMIEYLGV	682
Db	707	stewyfvvgjfoalnglqpaftslfslrslgftfcdedqpeckrgnsmfslfvlvgi	766
QY	683	ICPFSYFPMQGLFYGRAGEIILTRMLRHIAFRKAMLYKODIAMPDEKNSGGTITLADIQ	742
Db	767	isllfllfgltfkgagellkrltrlyvtsmlrldqswtdobpknltgalrtllandaag	826
QY	743	IOGATGSRIGVLTQNAATFMGLSVTISFTYEMEMTFLLSIAPVLAVTGMETFAMTGFAN	802
Db	827	vkgaigrslavltqtnanlvgllslsllymqtlclllalvplialaayvcmkmslsgal	886
QY	803	KDKOELHAGKILATLEALENITRTVLSLREKAFEDQEMEMLOTOHRNTSKAQLTSGCYAF	862
Db	887	kdkkelegagkrlatealenfrtvtvslcreqkfeymyqslqpyrnsrltkahlfyvsfsi	946
QY	863	SHAFITFYFAAGFRFGATLLIOGMMTPEBGMFIYPTALAVGAMALIGKTVILAPYSKAKSG	922
Db	947	lqammyisyagcftfigalvlvaneftmfgdvlvlvlsalvlgamavggvssfapolyakavs	1006
QY	923	AAHLFALLERKPNIDSRSOEGKKPDCTEGNLEFREVSFFPCRPDVFILKGLSLSTIERGK	982
Db	1007	aahvltmiekeprlidsysphglkpnltlegnvtfnfevfyuptrpdpdlvglsalevkkqg	1066
QY	983	TVAFYVGSSGCKSTSYVOLLRPLRDPVOGQVLPEDGVDAKELNVOMLSQALIVYQBEVLN	1042
Db	1067	tlalygssgcgkstevvqllderfyqplagsvllldgkelkhlvnylahlglyvsqepdlld	1126
QY	1043	CSIAINIAVGNNSRVPRLDERDEIKEAANANIHSHFIEGPEKYNTQVKGKAGLSGGGQORL	1102
Db	1127	csiaenladygnstrvsineelvgakaealnhihtetlpekyntcrvgdkgqsls9gqkgt	1186
QY	1103	AIARALQPKILLLENTSALDNDSERKVVQVHALDKARTGRFCLVTHRLSAIQNDLIY	1162
Db	1187	alatarlvqrphlllldestsaldteseklvvgyealdkaregrtcclvahrlstlqnadli	1246
QY	1163	VLIHNGRIKEOGTHOELLRNDRIVKLVNAOS	1193
Db	1247	vftngkvkvehgtbhqllaqkgilymsvsqg	1277
RESULT 2			
AAAB81065	ID		
AAAB81065	standard; Protein: 1283 AA.		
AAAB81065;	AC		
AAAB81065;	AA		
25-JUN-2001	(first entry)		
Cynomologous monkey P-glycoprotein variant 2.	DE		
Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRL;	XX		
efflux pump.	KW		
Macaca fascicularis.	OS		
Location/Qualifiers	XX		
Misc-difference 93..95	FH		
/note= "An additional 3 amino acids are present compared	FT		
to PGP variant AAB81064"	FT		
WO200123565-A1.	XX		
05-APR-2001.	XX		



```

XX PD MO9740160-A1.
XX PF 30-OCR-1997.
XX PF 24-APR-1997; 97WO-N000216.
XX PR 24-APR-1996; 96EP-0201094.
XX PA (UYGR-) RIJXSUNIV GROOTINGEN.
XX P1 BoJhuis H, Konings WN, Van Veen HW, Venema K;
XX DR WPI: 1997-535844/49.
XX PR Prokaryotic homologue of human multiple drug resistance protein -
XX PF used to screen for compounds that inhibit, or avoid, drug resistance
XX PS Claim 10; Fig 1; 35pp; English.
CC CC The present invention describes a recombinant or isolated nucleic acid
CC CC (1), derived from a prokaryotic gene, which encodes at least a specific
CC CC and/or functional part of a transporter protein (7P), or its
CC CC derivatives, which has functional and/or structural similarity with the
CC CC P-glycoprotein (PG) encoded by the human multidrug resistance
CC CC (MDR)1 gene. The present sequence represents the human MDR1 protein,
CC CC derived from MDR1-N and MDR1-C as shown in the specification in
CC CC figure 1. (1) is used to express recombinant proteins; its fragments
CC CC are also useful as probes and primers for detection and amplification
CC CC of related DNA. The protein produced, or cells expressing them, are
CC CC used to determine if substances can inhibit, or avoid, MDR proteins,
CC CC and in a screening method for identifying compounds that inhibit
CC CC transport of cytotoxic substances from cells. Also, cells with a
CC CC transmembrane protein, especially where expressed from (1), can provide
CC CC (additional) MDR, particularly for use as a model system to study
CC CC mechanisms of action of PG.
S0 Sequence 1280 AA;

Query Match 59.1%; Score 3554; DB 18; Length 1280;
Best Local Similarity 55.6%; Pred. No. 4.1e-293;
Matches 682; Conservative 246; Mismatches 263; Indels 36; Gaps

OY 1 MILGLASIVNGACPLMPVLVGEUSD-----NLISGLVOT--NTYSEF----- 43
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 51 mvvgllaiaiiingagp|lmllvfgemt|d|faanagled|msnt|sntsd|ndtgf|fmleed 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 44 --RLTLYVVGICVALITFGYIOISLWITPAAQTRIRKOPFHSVAADIGFSDSDIE 101
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 111 mlttyyyysg|yagv|vaay|yqstwc|laagq|lhtk|tq|fhalmr|ge|yfd|ndv|ye 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 102 LNTFMT-DIDKISDIDGKIALLFQNMSTFSGIAGLVKYGKLLVLTSTPLMASAA 160
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 171 lntrltdvskine|g|dk|gm|f|fgsmat|ft|fg|iv|ftrgk|k|t|lv|a|s|p|v|g|saa 230
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 161 ACSRAVITSSELSA|YSKAGAVAEV|SISRTV|A|FRAQEKELORYTONLKDADGFK 220
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 231 vwakl|ss|ft|cd|k|l|aya|kag|a|eev|la|a|t|t|v|a|f|gq|k|e|l|e|y|n|l|e|a|k|r|g|t|k 290
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 221 RTIAKVSIGAVYFPMNGYGLAFMYG|SL|INGE|PGY|IG|V|LVA|F|E|V|H|SS|V|C|IG|A 280
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 231 kaitn|st|ga|f|l|l|y|s|a|l|a|t|a|y|g|t|l|v|s|g|e--y|s|g|v|l|t|f|s|v|l|g|a|t|s|v|g|a 348
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 281 VPHE|T|F|A|I|A|R|G|A|P|H|I|Q|V|I|D|K|R|S|I|N|F|S|T|A|G|K|P|E|S|I|E|G|T|V|E|K|N|S|F|N|P|S|R|S|I|K 340
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 349 sps|lea|f|a|n|a|r|g|a|e|y|e|f|f|l|d|h|k|p|s|d|s|y|s|k|g|h|k|p|h|n|k|n|l|e|f|r|n|h|f|y|p|s|t|k|e|y|k 408
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 341 ILK|G|N|L|R|K|S|E|T|V|A|L|G|L|S|G|S|T|V|Q|L|L|Q|L|Y|D|P|D|D|G|F|I|W|D|E|N|D|I|R|A|L|N|V|R|H|Y|D 400
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 409 ilk|g|n|l|k|s|g|e|t|v|a|l|g|l|s|g|s|t|v|q|l|l|q|l|y|d|p|d|d|g|f|i|w|d|e|n|d|i|r|a|l|n|v|r|h|y|d 468
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 401 HIGV|S|Q|E|P|V|L|F|G|T|I|S|N|N|I|K|G|R|D|V|D|E|E|R|A|R|E|N|A|Y|D|I|F|E|P|N|K|R|N|T|L|V|G|E|G 460
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Dh	469	liyvsgsepvlfactlaenlrygreuntdelckavkeanaydfimlkhphkfdtlvsgerg	528
Qy	461	AKMSGGQKQRIATARALVNRPKLILDEATSAIDSEKSAVQAALEKSKGRTTVVAHR	520
Dh	529	aqlesgqgqrlataraIavrnprklllIdeatsaldeseavvvaldkarkgfttlviah	588
Qy	521	LSTRASDLIVYTLKDGMLAKGAHAELMARGLYSLVMSQ-----DIKRDQEGMEWT	574
Dh	569	IstvrnadvlagfdgglyvekgndelmekkglyfkIvlmqagaveleenaadesksei	648
Qy	575	YSTERKTSPLPSHV-----KSIK-SDFDKAEESTQKSEISPEVSLIKTLKLNKPEM	627
Dh	649	dalemsndsrssliirstrcrsvrgsqadnrlstlkealdeslprvstwrImklnlcw	708
Qy	628	PFVVLGLTSLVNLGTVHVFVESILFAKTIIMEGN-NDKTKLKHDAEYISMFVILGICV	686
Dh	709	pyfvaygfcalingqlpafalifskliqyftcidddpektqrnsllfsalfalaglsfifi	768
Qy	667	SYPMQGLFYGAGGILLMRRLRHAFKAMLYODIAMDEKENSQGLTTLTADINOIGA	746
Dh	769	ftlfgftfIgaageIklkrltymfismrgdovswddpbnrttgalctrlandaagvka	828
Qy	747	TGSRIGVLTQATMNGSLVSIISFYIGWEMFTLISAPLAVATGMIETAAMGFANKQK	806
Dh	829	IgsrlavltqnalalqglisfsfygqlllIlalvplatalagvemmIsbgqlkdk	888
Qy	807	ELKHAGKIALEALENIRTVISLTREKAEQOMEYEMLOTQHRNTSKKAQIISCVAFSHAF	866
Dh	889	elegagkIatealenftvtvsltdqekfemnyqagslqpyrnsIklahfigltfstgam	948
Qy	867	IYFVYAGFRFGATLIDQAGMTPBGMIVYTTALAGAMALIGKTVLAVPYSAKSGAHL	926
Dh	949	myfIvagqftfIgaIvabhImstfcdvllvtsavfIgamavvgvgsfapdyakakIsaahI	1008
Qy	927	FALLERKPNIDSRQOEKRPDCEGNLEFREVSFFYPCRPDVFILNGLSLSIERGTYAF	986
Dh	1009	ImietekrpldsystegImptnlegnvtfgvfnfprcpdIprvglsIekvkgqtlal	1066
Qy	987	VGSSGCGKTSVOLLRLYDPVOGQVLFEDVDAKELNVQMSQIAIYQEPVLEFNCISA	1046
Dh	1069	vgssgcgkstvtvqlllerfpyrlagkvllldgkelkrInvgqlvrahlgIvsgperlIdcsia	1128
Qy	1047	ENIAYGDSKRVPLDEIETKEAANAANHSFIEGLPEKRTNOVGLKQOLSGGQKQIATAR	1106
Dh	1129	enIaygdnsIrvsgsealviraakeanIhaflesIpnkystekvgdkqfqlsggqgqrlatar	1168
Qy	1107	ALLQPKLILDEATSAIDSEKSAVQAALEKSKGRTTVVAHR	1166
Dh	1189	alvtrpnhlllIdeatsaldesekevvaqaldkaregrtctviahnrlstIqnadllvtfqn	1248
Qy	1167	GRIKEQGTQHOELLRNRIYFKLYNAOS	1193
Dh	1249	grvkehgtlqglllaqkgIysmvsvga	1275
RESULT 4			
ID	AAV58186	standard; Protein; 1280 AA.	
XX	AAV58186;		
XX	14-MAR-2000	(first entry)	
DE	Human wild-type multidrug resistance-1 (MDR-1) protein.		
KW	Multidrug resistance; MDR-1; P-glycoprotein;		
KW	transmembrane efflux pump; hematopoietic stem cell; transduction;		
KW	bone marrow transplantation; chemotherapy; radiation therapy; cancer;		
KW	gene therapy; gene replacement; genetic defect; thalassemia;		
KW	Gancher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;		
KX	cytokine; wild-type.		
OS	Homo sapiens.		

[illegible]

QY	221	RTLSKYSVLAVYFPMNGTIGLAFWGTGSLILNGEPCITIGIVLAFVSVYSHSXYCIGA	280
Db	291	ka1an1s1gaaf11yasaya1a1fwygtc1v1sge--'ys1agv1lv1f1v1s1g1a1fsvq4	348
QY	281	VPHETFA1AGAAFH1FOYIDKPPSIDNFP5AGYVPE5IECTVEFKNVSNYSRPSIK	340
Db	349	sp1s1lea1ana1gaayef1k1i0k1p1s1d1s1y1s1g1h1p1d1i1k1n1e1f1nv1h1s1y1s1p1a1r1k1evk	408
QY	341	ILKGLMRKISGETVALVGLNGSGKSTVYQLORLYDPDDGFIWDENDIRALVNRHXR	400
Db	409	1lk1g1n1k1v1s1g1y1q1v1a1l1y1g1n1s1g1c1k1t1v1q1m1g1r1y1d1p1e1g1m1v1s1v1d1g1d1i1t1v1f1r1e	468
QY	401	HIGVVSQEPVFGTITSNNTKRYGRDDVYDDEBNRARAENAAADFIEMEPNKNLTVGEKG	460
Db	469	11l1v1s1g1e1p1v1f1a1t1a1e1n1y1g1r1e1n1v1u1n1d1e1k1a1v1e1a1n1a1d1f1m1k1p1h1f1d1t1v1g1e1r1g	528
QY	461	AQMSGGKORAIARALVNRPKILIIDEATSAIDSEKSAVOAALEKASKRTTIVAAH	520
Db	529	aq1s1g1g1q1k1r1a1t1a1r1a1l1v1n1r1p1k11111d1e1a1t1s1a1d1e1s1e1a1v1g1v1a1d1k1b1r1t1c1t1v1a1h1r	588
QY	521	LS1RSADLVITLKGMLAEKGAHAELMAKRGLYSLVNSQ-----DIKKADEQMSMT	574
Db	589	1s1t1v1n1a1d1v1a1g1f1d1d1y1a1v1e1k1g1n1d1e1l1m1k1e1j1y1k1f1v1e1m1q1a1g1n1e1v1e1n1a1e1d1e1s1k1e1l	648
QY	575	YSTERKINSLPHV-----KSK-SDFIDKAEESTQKESLSEPVSLKILKLNKPEW	627
Db	649	d1a1e1m1s1n1d1e1s1r1s111k1r1s1t1r1i1v1s1y1s1g1q1d1k1s1t1s1k1e1a1d1e1s1p1r1v1s1f1w1m1k1n1l1e1w	708
QY	628	PEVVLGTASVLNCTVAPVPSIIIFAKITMFCN-NKTKTLKDAEITVMIFVILGVCFV	686
Db	709	p1y1v1v1y1f1c1a1i1n1g1l1q1p1a1f1a1i1f1s1k1i1y1f1t1r1d1d1e1p1e1k1r1n1s1f1l1f1a1g1i1s1f1i	768
QY	687	SYMQLGFYGRAGEITLMRLRHLAFKAMLYODIAMEFDEKENSTGGLTTLAIDIAIQGA	746
Db	769	t1f1f1q1g1f1f1g1a1g111k1t1r1y1m1f1s1m1r1q1d1v1s1w1f1d1q1b1k1n1t1c1g1a1t1r1l1a1n1d1a1a1y1k1a	828
QY	747	TGSRGVLTQNAATMMGLSVIISIFIVEMEMTFLISLAPLAATGMIEAATMGAFKANDQ	806
Db	829	1q1s1t1v1a1l1t1q1i1a1n1d1l1g1q1i1s1f1i1s1y1w1q111111a1l1v1p1l1a1a1y1a1v1e1m1k1s1g1q1a1l1d1k1	888
QY	807	ELKHAGK1ATEALEN1NTIYSLTRKFAFQDMYEEM1QTOHRNTSSKKAQIISCYAFSHAF	866
Db	889	e1e1g1a1g1k1a1t1e1a1e1n1f1t1v1s1l1c1q1e1k1f1e1m1y1a1q1s1p1y1r1n1s1l1k1a1h1f1g1t1s1f1q1a1m	948
QY	867	IYFAVAAGFREGAYLIOAGNMPDEGMFYFTALATAGAMAIGKTLVLADEYSKANSGAHL	926
Db	949	m1y1s1a1y1g1f1r1f1g1a1y11v1a1h1k1i1m1s1f1e1d111v1f1s1a1v1f1g1a1m1a1v1g1v1s1f1a1p1d1y1a1k1a1s1a1h1	1008
QY	927	FALLEKRPNDISRSQBEKKRDDCEGNLEPPREVSFFYPCHQDDVFTIRGISLISIEKGATAF	986
Db	1009	1i1m1l1e1k1r1p1l1d1s1y1t1e1g1i1m1p1n1c1l1e1g1n1t1e1g1e1v1f1n1p1t1r1d1p1r1v1q1i1s1l1e1k1y1q1t1a1	1068
QY	987	VGSSGGCGKSTVOLQRLYLPVQGOVLFPGVDAKELNVOMLSQIAIYPOEPVLFNGSIA	1046
Db	1069	v1g1s1s1g1c1k1s1t1v1q1l1e1f1r1y1p1a1g1k1v11d1g1k1e1k1r1n1v1g1w1l1a1h1g1y1v1s1g1p1l1f1d1e1s1a1	1128
QY	1047	ENIAYGDSNRVPLDEIKEAANANHSFIEGLPEKXTNQVLKGAQSLSGQOKORLAIR	1106
Db	1129	e1n1a1y1g1d1s1r1v1s1q1e1i1v1r1a1e1k1a1n1h1a1f1e1s1p1n1k1y1s1t1k1y1d1k1t1q1s1g1q1k1q1t1a1r	1188
QY	1107	ALLQPKLILIDEMTSLNDMSRQVOMHLDKRRGRCLVYTHHLSATONADILVYLN	1166
Db	1189	a1l1v1g1p1h1l1l1d1e1a1t1s1a1d1e1s1e1v1v1g1e1a1d1k1a1r1e1g1t1c1t1v1a1h1s1t1c1n1a1d1l1v1n1f1q1n	1248
QY	1167	GKIEQGTQHOELLNRNDIYFKLVNAOS 1193	
Db	1249	g1r1v1k1e1h1g1h1q1l1a1q1k1g1f1s1m1s1v1q4 1275	

XX AAB81959;  
AC  
DT 02-JUL-2001 (first entry)  
XX  
DE Human MDR1.  
XX  
KW Human; MDR1; multi-drug resistance; cholesterol uptake;  
KW hypercholesterolaemia; hypocholesterolaemia; atherosclerosis;  
KW coronary artery disease; cerebral vascular disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200121762-A2.  
XX  
PD 29-MAR-2001.  
XX  
PF 22-SEP-2000; 2000WO-US26099.  
XX  
PR 23-SEP-1999; 99US-0155819.  
XX  
PA (BARN-) BARNES-JEWISH HOSPITAL.  
XX  
PI Stenson WF, Tessner T;  
XX  
DR WPI; 2001-328100/34.  
XX  
PT Modulating cellular cholesterol uptake for treating hyper or  
PT hypocholesterolemia, involves administering an agent that inhibits or  
PT increases the expression of multiple drug resistance-1  
XX  
PS Example 1; Page 43-47; 47pp; English.  
XX  
CC The present invention describes a method of modulating cholesterol uptake  
CC in cells, involving administering an agent capable of inhibiting the  
CC multi-drug resistance protein MDR1. This is useful in the prevention and  
CC treatment of cholesterol-related diseases, including  
CC hypercholesterolaemia, hypocholesterolaemia, atherosclerosis, coronary  
CC artery disease and cerebral vascular diseases. The present sequence is  
CC the human MDR1 protein.  
XX  
SO Sequence 1280 AA;  
  
Query Match 59.1%; Score 3554; DB 22; Length 1280;  
Best Local Similarity 55.6%; Pred. No. 4.1e-293;  
Matches 682; Conservative 246; Mismatches 263; Indels 36; Gaps 9;  
  
QY 1 MILGILASLVNGACLPMLPVLVGENSD-----NLISGCLVQT--NTYSFF----- 43  
DB 51 mvvglaaahhagipmlmlvfgemtdifanagnledlmsnltmrdsndtgcffmnlced 110  
QY 44 --RLTLVVGVALILGCTIQISWITTAARQTKRIKQFPHSHVLADIDIGFDCSDIGE 101  
DB 111 mtrayayysgigaylvaylqvstwciaaqrqkirkqfthamreigvfdhdvge 170  
QY 102 LNTKMT-DIDKISDIDGKIALLFQNMSTFSGIAVLGVKGLKTLVLTSLTSMASAA 160  
DB 171 lntitddvsktneigqkigmfsgmatffgftvgftrgkltvlalaspvlglsaa 230  
QY 161 ACSRWVILTSKELSAISKAGAAVEVLSTRTVIAFRAQEKELQRYQNTLKADKQFQIK 220  
DB 221 vvaklissftfdkellayakagavaeevlaaitvtafgqkkeleryknlleaakriglk 290  
QY 221 RTIASKVLGAVYFFMNMNTYGAFWYGTSLIINGEGTIGVLAVFPSVHSSICIGAA 280  
DB 291 kaitenaisigaflilysayalatygtlvisge--yslqvltvifsvlgaitsvga 348  
QY 281 VPHFEFTFARGAAPHIQVIDKRPISINFSTAGYKPSIEGTVEFKMVSFNPSPSIX 340  
DB 349 spstaefanaragaeyefkikidnksidsyskghpnidkynlefrnhfsysprkewk 408  
QY 341 ILKGLNLRKISGETVALVGLANGSKSTVVOQLQRLYDPDGEFIMVDENDIRALNVRHYRD 400

DB 409 ilkglnlkvgsgqctvalvngscgkstcvqimqrlydpdegmvsvdgqdlitlvrtfcre 468  
QY 401 HIGVVOEPVLEFGTTISNNIKYGRDQVYDEEMERAAERANAYDTIMEPNKFNLTAVGEC 460  
DB 469 liyvsgqevlftattiaenliyygreenvmdetkavkeanaaydfimkphkfdclvgerg 528  
QY 461 AOMSGGOKRIARALVARNPKIILDEATSAIDSESKASVQAALAKASKRTTIVAAHR 520  
DB 529 aqlsggqgqiaialaralyrnpkllldeatsaldtseavvqaldkarkgrttivlahr 588  
QY 521 LSTRSADLIYTLKDGMAELKGAELMAKRGLYSLVMSQ-----DIKADQMSMT 574  
DB 589 lstrnadvlagfdqyvlvekgndelmekgyifklytmgtagnevelenaadesksel 648  
QY 575 YSTERKTSPLHSV-----KSIR-SDFIDKAESSTQSEKISLPEVSLITLKNPEW 627  
DB 649 dalemsndstrssalrkstrsvrsgaqdrkstkadeislpvpsfwrkmlnltew 708  
QY 628 PEVVLGTLASVLTNGTVPFVPSIIRAKITMFGN-NDKTTLKHDAEISMVILGVICFV 686  
DB 709 pyfvvgvcalinglqgafalifsklgyvtrtdpctrqnsnlfelltalglisfl 768  
QY 687 SYFMQGLFYGRAGETILMKRLHAFKAMLYODIAMFDEKENSSTGGLTTIADIAQIGA 746  
DB 769 tfllqgftfgkagelltkrltymwfrsmrlrgvswfdpdkntgaltrlandaaqykga 828  
QY 747 TGSRGVLTQNAITMNGLSVSIISFTYGMETFLIISIAVLVGTQMEATFANGDKO 806  
DB 829 lgsrlavtqqlanlgvqlilslfgywqlcllilaipvialagavemkmslsgalokkk 888  
QY 807 ELKHAGKIATALENIRITVSLTREKAFQMYEEMLTQHRNTSKKAOIISCYAFSHAF 866  
DB 889 elegagkiatealenfrtvalstqeqkfemhayslqypynslrkaahlfiftsfqam 948  
QY 867 IYFAYACFPFGATYILQGRMTPBGMFVFPALAYGAMATKTLVLAPEYKASGAAHL 926  
DB 949 myfsyagcfrtgyavahklmsfedvllvfaavfygamaqvqsfpdyakaklsaaht 1008  
QY 927 FALLEKKPNIDSRSGEKKPDTCEGNLEFRVSPFPCRPVFLRGSLIEGKTYAF 986  
DB 1009 imllekrplidsystcegimprntlegvnltgevinhyrtpdipvlgalslwkqgqltal 1068  
QY 987 VGSSGCGKSTGVOLQRLYDVOGVLEFDGVDAKELANQMLRSQAIYPOBVLFNCSIA 1046  
DB 1069 vgssegcstvvqlletfydpdlaqvlldkjkrlvngvtrahglvsgpdlfdcsia 1128  
QY 1047 ENIAYGDNRRVPLDEIKEAANANISFISGLPEKTYOYGLGAQISGOKORLAIR 1106  
DB 1129 eniaygdnsrvsgqeivraakeanlhaflieslpnkysctkygdkyqtqisgqkqialar 1188  
QY 1107 ALLOKPKTILLDEATSAIDNSEKVVQHAALDKARFGCLVWTRHLSAIONADLIYVLN 1166  
DB 1189 alvtrpnhllldeatsaldtsekevvgaalakaregrtciviahrlstclgnadliivfqn 1248  
QY 1167 GKIKEGTTHOELLRNNDIYFKLVNAOS 1193  
DB 1249 grvkeghyhgqlaqkgylyfsmvsvga 1275  
  
RESULT 6  
AAED0308  
ID AAED0308 standard; Protein: 1281 AA.  
XX  
AC AAED0308;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Dog P-glycoprotein (Pgp) allelic variant (genotype A).  
XX  
KW Dog; P-glycoprotein allelic variant; Pgp; multidrug transporter;  
KW MDR1; drug bioavailability; transgenic animal; genetic model.  
XX

OS Canis familiaris.  
 XX Key Location/Qualifiers  
 XX Misc-difference 197  
 FT /note= "His of Genotypec substituted by Gln"  
 XX  
 XX MO200123540-A2.  
 XX  
 XX PD 05-APR-2001.  
 XX  
 XX PF 28-SEP-2000; 2000WO-US26767.  
 XX  
 XX PR 28-SEP-1999; 99US-0156510.  
 XX  
 XX PA (GENT-) GENTEST CORP.  
 XX  
 XX PI Stocker PJ, Steimel-crespi DR, Crespi CL, Relf TC, Patten CJ;  
 XX WPI: 2001-235373/24.  
 DR N-PSDB: AAD03504.  
 XX  
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX  
 PS Claim 17: Page 91-93; 11pp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterized by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant.  
 CC This sequence is also referred as Genotype A protein. The PGP  
 CC enzyme functions as an efflux pump exporting small molecules  
 CC across the cell membrane. This enzyme is a member of the ABC  
 CC transporter family.  
 CC  
 CC Sequence 1281 AA;  
 SQ

Query Match 59.1%; Score 3554; DB 22; Length 1281;  
 Best Local Similarity 55.4%; Pred. No. 4,1e-293;  
 Matches 682; Conservative 238; Mismatches 269; Indels 42; Gaps 8;

QY 1 MTLGLASLVNACPLMPVLGEMSDNLISGCL-----VQNTYSPFR--- 44  
 DB 51 mlvgmaeiibhaalpmlmlvfgmntdfanagisrntkfpvlinesinnqghfinhle 110  
 QY 45 -----LTLVYVIGVAAALFEGYIOTISLWITTAAROTKRIRKOPFHSVLAADIGMEDSCDI 99  
 DB 111 eemtyayyysgigagvivaayiqvswfclaaqrqlkirkqfhnaimqegwfgdvhdv 170  
 QY 100 GELNTRMT-DIDKISDGGDKTALLFQNMSTFSIGLAVGLVGMKTLVTLTSLPLIMAS 158  
 DB 171 gelntrltdvavkinegldgkigmfqsatftfvgftrgwklvllaipvlgis 230  
 QY 159 AAACRMVISTLSKELSAVSKAGAAVEEVLISRTVIAFRAOKEIQRTOYLKDAKDFG 218  
 DB 223 aaiwakllsftckelleyakagavaeevlaairtviafgqgkkelerynkhleekyig 290  
 QY 219 IKRTTASVSLGAVTFPMNGTGLAFWGTSLILNGEPGYTIGTVLAVFVSIIHSSYICIG 278  
 DB 291 lkkaltanistgaaflliyasayalafwytsivlsae--ysigvltvfsvfsligafsig 348

QY 279 AAVPHETFAIRGAAFHIFQVYIDKKPSIDNEFTAGYKPEISIEGTVEFKNVSEFNPSRPS 338  
 DB 349 gaspsieafanargaaeyelfkldnkpsidsysksghkpnknlkgnelfnnvfispsrhe 408  
 QY 339 IKILGNLIRIKSGETVALVGLNGSGSTVOLLQRLYPDDEGFIWVDENDIRALNVRHY 398  
 DB 409 vkllkglnlkvgsgqvalvngsgcgkstctvglmqrlpdcgmcidgqdlrtlnvrhl 468  
 QY 399 KDHIGVSOEPVLFEGTTISNNIKYGRDDVYTDDEMEAAAEANAYDPIFMPKKNFTLVGE 458  
 DB 469 reitgvsqepvlfaltlaeniryrenvtmdelkavkeanaaydlm1pnhfdllyge 528  
 QY 459 KGAOMSGGOKRATARAIVRNPKIILDEATSAIDSEKSAVQAALAKASGRRTIYA 518  
 DB 529 rgaqls9gqkqrlaalaravrnphk1lldeasaldeseavvqaldkarqrrtllya 588  
 QY 519 HRLSTIRSAIDLVTTLKDGMLAEKGAHAEAMAKGLYSLVMSODIKKAODEMSMTYTE 578  
 DB 589 hrlstvrnadvlaqfdgdyivekgnhdelmkelyfklvltmg-trgnelelenatgesk 647  
 QY 579 KKTNSLPLHVSXSIKSDFI-----DKAESSTQSKELSLDEVSILKTLKN 623  
 DB 648 sesdalem-spkdsqssllkrrstsrshapqgdtklgkedlneavpvpsfwrllkin 706  
 QY 624 KPMPFVVLGTLASVNLGTVHPFESIIFAKITTF-GNNDKOTLKDAEYSIMFVLGV 682  
 DB 707 stemprfvvvglfcalingdlpafslisrlilgltredpeptrkgnmtvylfvlgi 766  
 QY 683 ICEVSYFMQGLFVYGRAGELITMLRLHAEKAMLYODIAMFDEKNSGTGLTTLAIDAO 742  
 DB 767 isftfflfgfctfgkegellkrlrymvfrsmrlqdvswfdpdkntglaltcrlandaag 826  
 QY 743 ICGATGSRIGVLTQNTNMGSLVYISFYIGWEMTFILISIAVPLANTGMIEFRAAMGFAN 802  
 DB 827 vqgaqsrilavltqnlantlgtllsllywqgl1llla1vllaagvcmkmls9gal 886  
 QY 803 KKOELKHGKATLEAKENIRITVSLTRKAFOMEEMJOHRTMSKAOIIGSCYAF 862  
 DB 887 kkkkelegagklatealenfrtvsiltreqkfeymagaglypyrmlskahlfyfsfsl 946  
 QY 863 SHAFFYAFAYAGFRGAVYLIOAGRMTPEGMFVFTAIYAGVAAIGTLYLABEYSKAKSG 922  
 DB 947 tqammyfsyagctrfgaylvaneffmfgdvlvfasivgamavgvssfapdyakaks 1006  
 QY 923 AAHLFALLEKKNIDSRSGEKKPDTCBGNLEFREVSFPYPCRPDYFIRGLSLSTERCK 982  
 DB 1007 aahvmlilekspildysphgikpnleqvntnevfnyptpdpilvlglslevkkg 1066  
 QY 983 TYAFVSSGCGKSTSVOLLQRLYPDQGOVLPFGVNAKELNOMWLSQAIAVPQEVLPN 1042  
 DB 1067 ctalvsgsgcgsstvvqllertfydplagsvlldgkaiikhnvqwrabrhgivsgepdltd 1126  
 QY 1043 CSIAENIAYGDNRRVPLDEIREKANAANINHSFIEGLSPKRYMTOVGLKAGOLSGCOKRL 1102  
 DB 1127 csaenaiaygdnrsrvshvshelngaaekanhfielelpekyntrvgdkqtqsggkqrl 1186  
 QY 1103 A1ARALQPKLILDEATSAIDNDESEKVVQNAHDAKARGRICLVVTHRLSAIONADLV 1162  
 DB 1187 a1aralvtrqphllllldeatsalidesekevgaaldkaregtcoivlahrlstlgnadliv 1246  
 QY 1163 VLHNGKIRKEOGTHOELLRNDRDYFKLVNOS 1193  
 DB 1247 vifngkvkqhngqllaqgilylfsmvsvga 1277

RESULT 7  
 AAE00309 standard. Protein: 1281 AA.  
 ID AAE00309;  
 XX AAE00309;  
 AC AAE00309;  
 XX AAE00309;  
 DT 13-JUN-2001 (first entry)



[illegible]

Db	231	aaiwaakllsftckellaykagaveeevlaairtylaagggkkelerynkhleakig	290
Qy	219	IKRTIASKVLSGAVYFFMNQTYGLAFWYGTSLILNCEPGYITGVLAFFSVYHSSYCIG	278
Db	291	lkkaitanisigaeflllyysayalaftwysgslisse--ysigvyltfvffsvlligaftsig	348
Qy	279	AAVHFETPALRGAAHITQVIDKKPSTIDNFTACVKKRESTIEGVFEKPNVSFNPSRPS	338
Db	349	qaspsiafaanaqaaaeftklidnkpstidsyskshkpdnklgmletknvftisypsrke	408
Qy	339	IKIKAGNLRIKSGEYVALVGLNGSSKSTVVOLOQLYPPDGFIMVENDIRALNRYH	398
Db	409	vkllkglnlkvgsqgtvalyngsgckstvtvqimqrllypbtgmcicdgqdrctlnvzhl	468
Qy	399	RDHIGVVSQEBVLFGTTISNNIKYGDHDTVEEMERAREANAYDFIMEFPKNTLVGE	458
Db	469	reltgvsvgepvlfatliaenitrygreynvumdeiekavkeanaqdfimlprkfdtlvge	528
Qy	459	KGAOMSGGQORIALAALVARNPKLILDEBATSALDNSEKSAVOALAEKASGRITTYA	518
Db	529	rgaqlsggqkyrtalalarnpkllldeatsalDNeseaavqvaldarqgrttlyla	588
Qy	519	HLRTISADLIYMLKMGMAEKGAHELMARKGLYSLVMSODIKKADQKQESMTYTE	578
Db	589	hrlstvmadylagfdggvlyevkxndelmekyglfxklyvtmq--trgnelelenatesk	647
Qy	579	RKTISSLPHSVKSIKSDFI-----DAAESTOSKEITSPEVSLKILKLN	623
Db	648	sesalalem--spkdsgrslkkrsttrrslhagqgdcklgtkedlnenpyvsefwrlkln	706
Qy	624	KPEPPEVVLGTLASVLNGTVHPVPSIIFAKIITMF--GNNDKTTLKHDAEITSMTFVLGV	682
Db	707	stewpyfvgvfcalnvglqpaftsllfsrlilgfrtredpeltkqysumfslvflvylgi	766
Qy	683	ICFYSYMOGLFGRAGEIITFMRHLHAFAMLYODIAFDEKKNSTGGLTTLILADIQ	742
Db	767	lsiflltqlgtltgkagelltkrltymvtrtsmlrgdvsfddpknltgaltrrlandaq	826
Qy	743	IOGATGSRIGVLONANMGLSVIISFYIWEWMTFFILISIAVLAVNGMIETAAMTGFAN	802
Db	827	vksglgerlavlqgnlanlgtjllsllysqwcltlllalyrlalagvemmkslsgal	886
Qy	803	KDKOELKHAGKIATLEALENIRITVLSLREKAFQOMEEMLQTOHRTSKKAOIISCTAF	862
Db	887	kdkkelagagklataeaenitrvysiltreqkfeymwagslqvpvynslrkanhlfvgsfsl	946
Qy	863	SHAFIYAVAVAGPEFGVLYLOAGMPEGEFIFYFTAIAGMAIGTLYLAPEYSKASG	922
Db	947	lqnamytsyagcrlfgyaylvanefmfgdvllyfsaivrgaanaqvqvsfapdyakavs	1006
Qy	923	AAHFALALEKKPNIDNSOGKCRPDICEGNLEBERSEFFYPCRPDVFILKGLSISIERK	982
Db	1007	aahrlymleksiportalidsyprbglkprntlegvntnevvfryprtdprdlvqslslevkqg	1066
Qy	983	TVAEVSOGCKSTSVOLQRLYDPVOGOVLFEDGVDAKELNVOMLRSQIAYVQEBVLFN	1042
Db	1067	tlavgsqcgkstevqlllerftrplagsvllqgkkelhvnqywlrahlngysvgerllfd	1126
Qy	1043	CSIAENIAYGDNKSVYVLDIETKEAANAANTHSTIEBELPERKYNTQVGLKGAQUSGQORL	1102
Db	1127	csiaenlaygdnsrvvsheelmakaeanhhfletlcpkynltrydgkyltqslsgqkqrll	1186
Qy	1103	AIATAALLQKPELILDEBATSALDNSEKVOYHMLDKAFARGCLVYTHTSIAONADLV	1162
Db	1187	aiataalvqrplllldeatsalDNesekeavqvaldarqgrtceivlahnlsicnadiiv	1246
Qy	1163	VLHNGKIKEOGTHOELLRNMDIFKVLVNAOS	1193
Db	1247	vfgngkvkhebtbqlllagqgylyfsmvsyga	1277

[illegible]

Db	230	lkaltatnaisigaeflliliasyalaftwgsfvlssse--ytlqgvltfvtfsvlqafalsg	347
QY	279	AAAPHEFEFALARAAAHPIHOVIDKRSIDNEFTAGCKPEPSTIGCTVEFNFNFNPSRS	338
Db	348	qaspsieafearaayaeifkildnksidsysksghkpnldkgnlefnvfnyspsstke	407
QY	339	IKIKGLMLRKISGEYALVGLNGSKSTVVDLQRLYDPDDGFIWDENDIRALNVRHY	398
Db	408	vkllkgnlkwgsgqvtalvagnscgkstvtlmgrltlypdrtdgmwvcdsgdrtlnvrhl	467
QY	339	RDHGVVSQSEPVLFGTITSNNIKRGDDVTDEMERARAPANAYPDIMEFPKFNPLYGE	458
Db	468	reltgvsvgepvltatlaenlirytgenvfmdelekavkeanaaydfimklnpkfdtlvyge	527
QY	459	KGAOMSGGOKRIRIARALVNPRTLLIDATSAIDSEKSAVOALKEASKGRRTIYVA	518
Db	528	rgaaisggqkqrtaataalvnrpkllldeatsalideseeavvqvaldkarxgrtltiya	587
QY	519	HRLSTIRSADILVTLLKQDMLAEKGAHMAELMAKRGLYVSLVMSODIRKADQESMTYTE	578
Db	588	hrlstvrnadvlagfdgdgvayekgnhdelmkexgyifklyltmq-trgnelelenatgesk	646
QY	579	RKTNSLPJHSAVSKISDPI-----DKAESGTSKSELSPEVSLKLTIKLN	623
Db	647	sesdalem-spksdgssslktrstlrtsihapqgqdkrlgtkedlnenpvsfwtlrlkn	705
QY	624	KPEMFVVLGLASLVGNVTPVPSIIIFAKITIMF-GNNDKTKLKDAEYIMFVILGV	682
Db	706	stewpytvgvltfcaalngglpapsllsftrtslfrtcdrepektrnsnmfsvlftvygi	765
QY	683	ICFVSYPMOGFEYGRAGEILTMRLRHIAFAMLYODIAMDEKENSTGGLTITTLAIDAO	742
Db	766	isftlftfgtftfkgelgelklrlryvnfsmrlqdwswtdobkxotgtlgtlrrlndaag	825
QY	743	IQGATGSRIGVLTQNAITNMGLSVITSTFYSMEWMTFLILSTAPVLAVTGMEIETPAAMTFAN	802
Db	826	vkgaigrslavltqnanlqglisllislygwgltllllaalvrlaagvwnkmkslsgal	885
QY	803	KDKQGLKAGKIAIALEALENTRTYSLRREKAPQOMCEMOTQHRMTSKKAOIISCAAF	862
Db	886	kdkkelagagklatealeenftlvslrtreqkfeymyqsglyqyrnslrtkahlfgvstsl	945
QY	863	SHAFYFPAAGFRRGATLIDAGRMEPEGMFIYFTALAYGAMAIGKTLVLAPEYSKASG	922
Db	946	tqammyfyagcftfgyavaneftmtdqtlvftsailvfgamavgvssfepdyakaks	1005
QY	923	AAHLFALLEKKPNIDSNSQEGKRPDTCEGNLEFREVSFFYPCRPDVFILRGSLSTERCK	982
Db	1006	aahvamlilekspilidsyphglkpnltlegnvtfnenvfnuprtprdiplqslsevkqg	1065
QY	983	TVAFPGSSGCKRSTSVOLQRLXPORGVOGLPFGVAKELNOMVLSQATIAPOEPEVFN	1044
Db	1066	tlavfsgsgcgkstvtvgllertfyoplagsvldqgkelklnvwlalrhagivsgdehlld	1122
QY	1043	CSIAENIYVGNNSRVPPLDEITKEKANAANAHSHFEGLPEKRYNVOVGLKGAOLSGGOKRL	1102
Db	1126	csiaenltygnnsrvshneelmgaakeanlnhffletpekyotrtrvgdkgtqlsggqkql	1185
QY	1103	ATARALLOKPKILLDEATSAIDNSKSVVOHALDKARIGRTGLVTVTHRLSAIONADLY	1166
Db	1186	atarelvtrpghllldeatsalidesekvvgaealdkaregtrctviahrlstlqnadliy	1244
QY	1163	VHNGKIEGCTHOELLRNRPYKVLNNAOS	1193
Db	1246	vftngkvkxngtchqlllaqkxyifysvsgva	1276
RESULT 9			
ID AAE00304 standard; Protein: 1280 AA.			
xx			
AC AAE00304;			

XX 13-JUN-2001 (first entry)  
 XX Dog P-glycoprotein (PGP) #2.  
 DE Dog; P-glycoprotein; PGP, multidrug transporter; MDR1;  
 KW drug bioavailability; transgenic animal; genetic model.  
 XX Canis familiaris.  
 OS  
 PN MO200123540-A2.  
 PD 05-APR-2001.  
 PF 28-SEP-2000; 2000WO-US26767.  
 PR 28-SEP-1999; 99US-0156510.  
 PA (GENT-) GENTEST CORP.  
 PI Stocker PJ, Steinmetz-crespi DT, Crespi CL, Relf TC, Patten CJ;  
 PI WPI: 2001-235373/24.  
 PS N-PSDB; AAD03489.  
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX Example 2; Page 72-75; 11pp; English.  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP). The  
 CC PGP enzyme functions as an efflux pump exporting small molecules  
 CC across the cell membrane. This enzyme is a member of the ABC  
 CC transporter family.  
 CC  
 CC Sequence 1280 AA:  
 50  
 Query Match 59.0%; Score 3549; DB 22; Length 1280;  
 Best Local Similarity 55.2%; Pred. No. 1,1e-292;  
 Matches 680; Conservative 239; Mismatches 270; Indels 42; Gaps 8;

DB 290 ikkalfanisiigaafillyasalatwylsvlse--ytlqvlvtffsvlgsatsig 347  
 279 AAVPFEFAIARGAAFHIFQVIDDKPESIDNFSTAGKPSIDGTVAEFKNVSENPSPRS 338  
 DB 348 gaapsiaefanagaayefikikdnkpsidsyeksghpnikgnlefkvnhtsyepke 407  
 339 IKILKGLNRIKSGEYVALVGLNGSGKSTYVQLQRLYDDDOGFIMVDENDIRALNRYH 398  
 DB 408 vkllkgnlkvsgqvalvagnsgcgsktvcqimqrllydpldvmwcdldgqdlitlnrhl 467  
 399 RDHGVVSOEPVLEGTISNNIKRGDVTDEEMERARANAYDFIMEPNKFNFTLVE 458  
 DB 468 reltgvsvgepvlfatiaenilryrenvmdelkavkeaneydfimkpnkfdtlvge 527  
 459 KGAOMSGOKORAIARALVNPRIILDEATSAIDSEKSAVOALEKASKGRTTIVA 518  
 DB 528 rgarlsggqqrilaralvrnpkllldeatsaldeseavvvaldkarkglttivia 587  
 519 HRUSTISADLIYTLKAGMLAEKGAHAEMLAKRGLYSLYMSQDIKADQMESMTSTE 578  
 DB 588 hrlstvrnadvlaglddgvlvckghdelmekgylfklvmq-rlgnetelenatgesk 646  
 579 RKTNSPLHVSYSKISDPFI-----DKAESOSKESISLPEVSLIKILKN 623  
 DB 647 sesdalem-spkdsqssllkrtrstrshapqgdrlgktdenenvsvsfrllkin 705  
 624 KPEWPFVVLGLIASLVNGTVHPVPSIIFAKITMF-GNNDKTLKHDAEITYSMFVLVG 682  
 DB 706 stwpyrfvvgficalinglqpaftsllfstrlgtfdepedekrqnsmtsvflvlg 765  
 683 ICEVSYFMGLFYGACELTMRHLAFKAMLQDIAMFDEKNSGGLTTLAIDIAQ 742  
 DB 766 isflftlgtftfkgaelikrlrlymfrsmllqdwswldpdrntgatlrandaaq 825  
 743 ICGATGSRIGVLRNANMGLSVITSPYGMEMFLLISAPVLAVMGMIETAMTGFAN 802  
 DB 826 vkqaglsravlcpnlanlqgllslislygqlclllalvpilaigvvekmalsqal 885  
 803 KDKOELHAGKIATELENTITVLSLREKAFEQMYEEMLOTOHRNTSKKAQITGSCYAF 862  
 DB 886 kdkkelagaglatcaenlftvsvlcreqfeymqslqvpynslrkhalfgvsfsl 945  
 863 SHATTFAYAGFRPGAYLIQAGKMTPEGMFTTATAVGMALGKTLVLAPESKAKSG 922  
 DB 946 tqammytsyagctifgylvanefmndvllvtsalvfgamavqyvsfapdyakav 1005  
 923 AAHLFALLEKKNIDSRQSEKKKPDCEGNLEFREVSEFFYPCRPDVILRGLSISIERGX 982  
 DB 1006 aahvymilekpslidsyphglkpnltleqvntfnevfnyftrpdlpylglslevkkg 1065  
 983 TVAFVSSGCGKSTSVOLLQRLYDPVQGVLPDGVDAKELNVQRLSOIATVPOEPVLEN 1042  
 DB 1066 tlalvsgcgskstvcqlerfydpplagsvldgkelhlnvqyralnlgvsgelplfd 1125  
 1043 CSTAENTAYGNSKRVPLDEITKEAANANHSFTEGPEKNTQVUGKAGLQSGOGORL 1102  
 DB 1126 csiaenlaygnsrsvshvehmqakeanlnhlfletlpekylntvrgdqbls99qqr 1185  
 1103 AIALRLQKPKRIILDEATSAIDNDSEKVOHAIADKARTGTCTLVYHRSIAONADLIY 1162  
 DB 1186 aialralvqphllldeatsldcesekvveagaldkaregtctlviahrlstlgnadli 1245  
 1163 VLHNGKIKEOGTHOELLRNRIYFKLVNAA 1193  
 DB 1246 vfgngkvehgtbqllaqkglfsmisvga 1276  
 RESULT 10  
 AAEE0303  
 ID AAEE0303 standard; Protein; 1281 AA.  
 XX  
 AC AAEE0303;

XX 13-JUN-2001 (first entry)  
DT  
XX  
DE Dog (Pgp) P-glycoprotein (genotype C) #1.  
KW Dog; P-glycoprotein; Pgp; multidrug transporter; MDR1;  
XX drug bioavailability; transgenic animal; genetic model.  
OS  
XX Canis familiaris.  
PN W0200123540-A2.  
XX  
XX 05-APR-2001.  
PD  
XX  
XX 28-SEP-2000; 2000MO-US26767.  
PE  
XX 28-SEP-1999; 990S-0156510.  
PR  
XX  
PA (GENT-) GENTEST CORP.  
PI  
XX Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
DR WPI; 2001-235373/24.  
DR N-PSDB; AAD03488.  
XX  
XX  
XX New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT Pgp inhibitors -  
XX  
XX  
XX Claim 17; Page 64-66; 11pp; English.  
PS  
XX  
CC The invention relates to dog P-glycoprotein (Pgp) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening Pgp inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by Pgp activity, by  
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids  
CC are used as oligonucleotide probes. Complements of Pgp nucleic  
CC acids are useful as antisense oligonucleotides, to induce a Pgp  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (Pgp) also referred  
CC as genotype C protein. The Pgp enzyme functions as an efflux  
CC pump exporting small molecules across the cell membrane. This  
CC enzyme is a member of the ABC transporter family.  
CC  
XX  
XX Sequence 1281 AA:  
SO

Query Match 59.0%; Score 3549; DB 22; Length 1281;  
Best Local Similarity 55.3%; Pred. No. 1.1e-292;  
Matches 681; Conservative 238; Mismatches 270; Indels 42; Gaps 8;

QY 1 MILDGLASLVAGALPLPLVLEKSDNLISGCL-----VQNTYFSFR-- 44  
DB 51 mlvgtmaailhgaalplmmllvfgmtcdsfanaglsrnktfvltnestlmtqhinle 110  
QY 45 -----LFLYVYGIVAAALIFGYIOISLIIITAAAROTKRIRKOFPSVLAODIGWDSQDI 99  
DB 111 eemltayayysgagvivaayiqvstfwaclaagrqilkrqfihnamrqegwtdvhdv 170  
QY 100 GELNTRMT-DIDKISDGIKIALLFQMSSTFISGLAVGLVKWKLTLVLTSTPLIMAS 158  
DB 171 gelntrltcdvskinegigdklqmfthstafctfgffvgrgwklrtvltaisprlqls 230  
QY 159 AAASRWVITLSKELSAKAGAAVEVLSIRIVTAFRAOEKLOQRTONLDAKDFG 218  
DB 231 aaiwakklstfdkellayakagaaveevlaaitrtvtaifsgqkelerynknleakqig 290  
QY 219 IKRTIASKVSIGAVYFEMNGTYGLAFWYGTSLINGEPGYITGVLAFFVSIHSSYICG 278

DB 291 ikkaltanisisgaafllllyasayalaftwgtstvlisse--yielgqlvtvfsvlgaftsig 348  
QY 279 AAVPHEFTFATARGAAHIFQYVIDAKKPSIDNESTAGYRPESIEGTVEKKNSFNPSPRS 338  
DB 349 gaspsieafanagaayalfkfidnkpsidsyskshkphdikhgldetknvhsfyspsrke 408  
QY 339 IKRLKGLNRIRKSGFTVALVGLNGSGKSTVVOQLRLDPDGGFMVNDENIRALNVBHX 398  
DB 409 vkllkgnlkwysgqvalvngsgcgskctlvqdmqrlypdcgdmvcidgqddrtlnvhl 468  
QY 399 RDHIGVVSQEPVLFGFTTISNNIKYGRDVTDEEMERAAEANAYDFIMEFPKNTLVGE 458  
DB 469 reitgvsqevplfatltiaenlrygrenvndelexakeanaaydfimklnpkftllyge 528  
QY 459 KGAQNSGQOKORIRARLVNRPKLLIIDEATSAIDSEKSAVQAALAKASKGRITTYVA 518  
DB 529 rgaqisgqkqrlaatalaralvnrpklllildeatsalideseeavvqalokarkgrttlvia 588  
QY 519 HRLSTIRSAADLIVTLKDGMLAEKGHAELMAKRGLYSLVMSODIKKADQESMTYS7E 578  
DB 569 hrlstvrnadvlaagfdgdvivekgnhdelmkkglyflvtmq-crigneletenatgesk 647  
QY 579 RKTNSLPHSVKSIKSDPT-----DKAESTOSKEISLPEVSLIKTLKLN 623  
DB 648 sesdalem-spksdssslkrtstrshabgqdrkigtcednenrpyvsfwrlkln 706  
QY 624 KPWEFVVLGTLASLVNCTVHPVESIIFAKITITMF-GNNDKTTLKHDAEIVSMIFVLGV 682  
DB 707 stewpyfvvgificalingltpafsfllsrlgltfrdedpckrqnsnmfsvlrlvyl 766  
QY 683 ICFVSYFVQGLFYGAGGILTMRLHLAFKAMLYODIAMFDEKESSTGGLTTLAIDIAQ 742  
DB 767 lsiflfflqgtftgkagelkrlkrymwfrsmrlrtdvswfdpkrktlgtaltrandaq 826  
QY 743 IQAGTSRIGVLTQATNMNGLSVIRSFYQEMPTLISIAPIVLAVTAMTIAATGRAN 802  
DB 827 vkgaigrslavltqgnialnglqllsilygwqltlllaivplaiagvnmkmsgal 886  
QY 803 KDKQELHAKIATEALENIRTVSLTREKAFEOMYEMLMQTOHNTSKAOKIIGCYAF 862  
DB 887 kdkkelegagklateatenfrtvalsrtreqfeynagslqvpyrnslrkahlifgvstfl 946  
QY 863 SHAFYFAVAGFRGATLQAGRMTPEGMFVFAIAYGAMAICKTLVLAPEYKASG 922  
DB 947 tqammyisyagcfrfagaylvanefmngdvllvisaiyfgamaqvqvsfapdyakakvs 1006  
QY 923 AAHLFALIEKKRPNIDRSQOEKKPDTCEGNLEFRVSFFYPCRPVFLIKGLSLIERGK 982  
DB 1007 aahvmitiekepildsypbhqikpntlegrvtfnevrtnypcrpdlpyrlqglstlevkqg 1066  
QY 983 TVAFVYSSGCGKSTSVQQLRLQYLPVQGVLEFDGDAKELNVQWLRSQIAIYPOEVLFN 1042  
DB 1067 tlalvgsscgskstvvgvlrlyfrydplagsvllidgfeikhvlnrahlgvayqsgplfd 1126  
QY 1043 CSIAENIAYGNSRVPYIDETKEAANAANIHSFIEGLREKRYTOVGLGAQLSGQOKRL 1102  
DB 1127 csiaenlaygdnsrvsvsneelmqaaeanlhfictlepkyltrvgdkygtqisgqkqr1 1186  
QY 1103 AIALRLQPKRLIDENSTALDNDESEKRVYQHALDKARFGKRVVTRRLSIAQVADLIV 1162  
DB 1187 alaralvtrpnhlllildeatsalidesekrvygealidkaregtctciavahrslrtgnadlly 1246  
QY 1163 VLHNGRIKQGTQHELNRNDYFKLVNAQS 1193  
DB 1247 vfqngkvkehgtbqllaqkgyifsmavsvga 1277

RESULT 11  
AAB81064  
ID AAB81064 standard; Protein: 1280 AA.  
XX  
AC AAB81064;



DE Human G185V mutant multidrug resistance-1 (MDR-1) protein.  
XX  
XX Multidrug resistance: MDR-1: P-glycoprotein;  
KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
KW gene therapy; gene replacement; genetic defect; thalassemia;  
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
KW cytokine; mutant; mutelin.  
XX  
XX Homo sapiens.  
OS Synthetic.  
FH  
FH Key Location/Qualifiers  
FT Misc-difference 185  
FT /note= "This residue is Gly in the wild-type MDR-1  
FT (AAV58186)"  
XX  
XX MO9961589-A2.  
PD 02-DEC-1999.  
XX  
XX 27-MAY-1999; 99WO-US11825.  
XX  
XX 28-MAY-1998; 98US-0086988.  
XX  
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
PI Sorrentino B, Bunting K;  
XX  
XX MPI; 2000-072615/06.  
DR N-PSDB; AAZ49333.  
XX  
XX Ex vivo expansion of hematopoietic stem cells transduced with a  
PT sequence encoding human multidrug resistance-1, used for bone marrow  
PT transplantation -  
XX  
XX Example 1; Page 82-90; 113pp; English.  
XX  
XX This sequence represents human G185V mutant multidrug resistance  
CC protein MDR-1, where the Gly residue at position 185  
CC of the wild-type protein (AAV58186) is replaced by Val. MDR-1 is a  
CC transmembrane efflux pump, responsible for the export of drugs from  
CC cells, particularly cancer cells. The wild-type MDR-1 shows increased  
CC resistance to etoposide and decreased resistance to vinca alkaloids  
CC compared with the G185V mutant. The invention relates to transducing  
CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified haematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC haematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from haematopoietic stem cells, e.g., thalassemia,  
CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in haematopoietic stem cells.  
CC Haematopoietic stem cells that express MDR-1 will be protected against  
CC chemotherapeutic agents, so can be engrafted while the patient is  
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
CC provides sufficient cells to permit standard biochemical analysis.  
CC Overexpression of MDR-1 allows cytokine-driven expansion of  
CC haematopoietic stem cells by at least 10-fold compared with a maximum  
CC of 4-fold in known procedures.  
XX  
XX Sequence 1280 AA;

Query Match 59.0%; Score 3545; DB 21; Length 1280;  
Best Local Similarity 55.5%; Pred. No. 2,4e-292;  
Matches 681; Conservative 246; Mismatches 264; Indels 36; Gaps 9;

QY 1 MTLGLIASLVNACLPMPVLYGEMSD-----NLISGLVQVT--NTYSFF----- 43  
DB 51 mvvgftlaahlgagiprlmmlvgemtldfanagnldmsnltntsrslndtqtfmfmleed 110

QY 44 --RLTLVYVIGVAALIFEGYIOISLIIITPAROTFRKIRKOPFHSVLAODIGMFSDIG 101  
DB 111 mrrlyvyyysgigaylvaaaylqvsfwclaaagrqhikrtqffhalmrgdygfdvhdvge 170  
QY 102 LNTRMPT-DIDKTSIDGDKIALIFQNMSTPFSIGLAVLGMKLLVTLVSTSPPLIMASA 160  
DB 171 lntrltdvskinevlgdklmgfifgsmatffcgivgftcrpykkltlvialspvlglsaa 230  
QY 161 ACSRWVISTSKELSAVSKAGVAEEVLSIRTVIAFRAOEKLEIQNTQNLDAODFGIK 220  
DB 231 vvakllsftcdkellayakagavaeevlaalrtvialfgygkkelerynlnleakrtlgik 290  
QY 221 RTIASKVSLGAAVYFPFMNGYGLAFWYGTSLIINGEPYIGTVLAFVFSVISHSYICGA 280  
DB 291 kaitanisigaaflilysayaalafygtltvlvsge--yslgvltvfsvllgafsvqga 348  
QY 281 VPHFTFALARGAAFIHFOVIDKKPSIDNFSTAGYKPESIECTVEKRVSNFVPSRPSIK 340  
DB 349 spsieafanargaayelrfkldnkpidsygsghkpdnlkgnlefrnvhfaypsrkev 408  
QY 341 ILKGLNLRKSGEYVALVGLNSGKSTVYVQLQRLYDPDPDGFIMDENDIRALNVRHYD 400  
DB 409 ilkglnlkvsgqvalvngsgcgskstlvqmqrlfydplegmvsyvgqgdrlrtlnvflre 468  
QY 401 HIGVVSQEPVLFGFTTSNNIKYGRDVTDEEMERARAEANAADFIMEFNKFTLWGEKG 460  
DB 469 ilgvvsgepvlftatltaenlrygreenvmdelckavkeanaajdfmlkphkfdtlvgerg 528  
QY 461 AQMSGQOKORIAIARALVBNPKILILDEATSALDSKSAVQAALEKASKGRTTVVAHR 520  
DB 529 aqlsggqkgrlatalaralvnpklllildeatsaldteseeavvaidkarkgtrtviahr 588  
QY 521 LSTISADILVTLKQGMALAEKGAHMAKRGILYSLVNQ-----DIKKADEQESMT 574  
DB 589 lstvlnadvlagfdgqvivekghndelmekgilyfkilvtagneveleenaadeeskel 648  
QY 575 YSTERKTNLSPLHSV-----KSIR-SDFIDKAESSTQSKESLPEVSLTKTLKLPKM 627  
DB 649 dalemsndsrslrlkrtrtsrvsgsqgdrtkskealdeislpvstwrlnklnltew 708  
QY 628 PNVVLGTLASVNGVIVHPVFSIIFAKITIMFGN-NDKTKLKHDAEILYSMIFVLGYICFV 686  
DB 709 pyfvvgvfaalngvqpfalfsklvgftriddpekrqnsnlfsllfalgylstfl 768  
QY 687 STFMGLFTYGRAGEILTMRLHATKAMLYODIAMDENKSTGCTTLTLLADIAOIQGA 746  
DB 769 tfflgqftgkageiltkrlymvtrsmirgqvswwfdpbkntgalttrlandaaqvkg 828  
QY 747 TGSRTGLVLTQNTMGLSVIISFIYGMWTFILISAPVLAVTGMETFAAMTGFANKDKO 806  
DB 829 lgsrtlaivtqnaianigtglisflfygwglllialvplialagvemmkslgqakdkk 888  
QY 807 ELKHAGKATPEALLENIRITVLSLREKAPDMYEMQLQTOHRTSKKAQIGSCYAFSHAF 866  
DB 889 elsegqklatealenfrtvsltqgekfemhmqasqvyprnslrkhalfgtlftstqgm 948  
QY 867 ITFATVAGRFAGYALIQAGRMPTBEGMFYVFTALATGAMAIGTIVLAPETSKAKSAALH 926  
DB 949 myfsyagctfrlgyavhklmsfedvllvfasvavfiamavgvssfadpakakisaah 1008  
QY 927 FALLEKKNPINDRSQEGKPPDCEGNLEFRVSFPVPCPDVFTLGLSLSTERGTVAF 986  
DB 1009 imliektplidsysteglmnpnltegnvltgevvfnyptrpdlvrglslvkwkqgtlial 1068  
QY 987 VSSSGGKSTVQLQRLPYVOGQVLFDDGVDAKELNVMQLRSQAIATBOEPLVFNCSIA 1046  
DB 1069 vsgsgcgketvqglarfydplagkylldgkeltkrlnvglrahlgivsegepdlifdcsta 1126  
QY 1047 ENIATGDNRRVPLDEIKANANANHSFIEGLPEKTYNTQVGLKGAQLSGGQOKORALATK 1106  
DB 1129 enlaygdnsrrvsvqeivraakeanthalteeslpnkysltkvgdkgtqsg9gkqrlatlar 1188  
QY 1107 ALLQPKLILLDEATGATSLDNSEKVVQHALDKARIGRTCLAVTHRLSAIQNDMLVILHN 1166





Db 1249 grvkehghqglldqkgyfsmvsvga 1275

## RESULT 14

ID AAE00306 standard; Protein; 1280 AA.

AAE00306;

13-JUN-2001 (first entry)

Human P-glycoprotein (PGP) #1.

Human; P-glycoprotein; PGP; multidrug transporter; MDR1; drug bioavailability; transgenic animal; genetic model.

Homo sapiens.

MO200123540-A2.

05-APR-2001.

28-SEP-2000; 2000WO-US26767.

28-SEP-1999; 99US-0156510.

(GENT-) GENTEST CORP.

Stockier PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;

WPI; 2001-235373/24.

New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful for determining the bioavailability of drugs and for screening for dog PGP inhibitors -

Claim 16; Page 78-80; 11pp; English.

The invention relates to dog P-glycoprotein (PGP) also referred as multidrug transporter (MDR1) and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog P-glycoprotein. PGP and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterized by PGP activity, by reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.

The present sequence is human P-glycoprotein (PGP). The PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.

Sequence 1280 AA;

Query Match 59.0%; Score 3545; DB 22; Length 1280; Best local similarity 55.5%; Pred. No. 2.4e-292; Mismatches 264; Indels 36; Gaps 9;

Matches 681; Conservative 246; Mismatches 264; Indels 36; Gaps 9;

1 MTLGLASLVNCAQLPLPLVLGMSD-----NLISGLVOT--NTYSFF----- 43

51 mvvgclaaibhagjplmmlvfgemtdlfanagnledlnsntnsdndtffmleed 110

44 --RLLYVGIVAAALRGYIOISLITTAQRTKIRKQFHFVSLADIGFNSCDICE 101

111 mtrlyayysgigayvlvaaylsvfcwclaaagrqfkhkirkfhaimgelsgyfdvdyge 170

102 LNTFRMT-DIDKISDGIKIALLPQNMSTFSGILAVGLVKGKRLTLVLTSTPLMASA 160

171 Intrltddvsklnewlgdklmgfmgmatffgflvgftrgkklvllalspvlglsea 230

161 ACSRWVSLTFSKELSAKAGVAEBVLSSIRTVIAFRAQEKELQRTYONLKDADFGIK 220

231 vvakllsftcdkllayakagavaeevlaaivrlvafggqkelerynkleeakrllgk 290

221 RTIASKVSLGAVYFMNGYVGLAFWYGTSLIINGEPGYTIGTVLAVFEFVHSSVCIGA 280

291 kaianisigaaflllyasayalaftwygtlvlsge--yslqglvlfvflsllgaitsvga 348

281 VPHFETFAIARGAFAHIFQVINDKPSIDNFTAGYKPESEIESTVEKKNVFNVPSPKSK 340

349 spsleafanaraaayefllidnkpseidsysgkphdnikgnlertnvhfyspskexk 408

341 ILKGLNLRKSGEVALVGLNGSKSTVVQLLQRLYDPDDGFLMDENDIRALNVRHYD 400

409 llkgnllkvsgqvtvalvngsgcgkstvtvqmqrllydpcgmwsvdgqdlrtlnvrflre 468

401 HIGVVSQEPVLFGTTISNNIKYGRDQVYDEEMERARERANADFTMEFPNKFMTWGEK 460

469 llgvsqepvllatlaenllygrenvundelekavkaanaaydlmkllphktdllvgerg 528

461 AQMSGGOKORIAIARALVNPKILILDEATSALDSESKSAVOAALEKASKGRTTVVAHR 520

529 aqlsggqkqrataralvnpkllldeatsaldeseavvayaldakqyrttlviahrr 588

521 ISTIRASDLIVTLKDGMLAEKGAHMAKRLYSLVNSQ-----DIKADQEMESMT 574

589 lstrnadvldfgddvlyvekgnhdelmekgyflklytmqclagneveleenaadeeskel 648

575 YSTERKTNLPLHSV-----KSIK-SDFIDKAEESTQSKELSLPVSLLKTLKLNKPPW 627

649 dalemsnmsrsllrkstrsvrgsqgdrklstkealideslprvstwrmlkllntew 708

628 PEVVLCTLASVNGVHPVPSIIFAKITPMFCGN-NDKTLKDAEIVSIFVLGYICFV 686

709 pyfvvgvfcainglqpfalfstkslyvfricddpekrqnsulfslflalaglstfi 768

769 tfflgftgkgealltkrllymvfrsmrlqdvswtdpdkntgaltrlandaaqyva 828

747 TGSRGVLTQNTMGLSVITISFYGEMTFLILSLAPLAVTGMETRAMGFANKDQ 806

829 lgsrlavtlqnanlgtllsflsywqltlllaivpilaagvemmllsgqalkdkk 888

807 ELKHAGKITEALDENRTIVSLTRKAFPEOMEYEMQTOHRMTSKKAQIIGSCYAFSHA 866

889 eilegagkisteatenfrtvsvlsgqekfemhysqslqyprynslrkhalfgltfstfgm 948

867 IYFAYAAGRFGAYLIQAGRMPEGNFIVETAIVAGAMAIGKTLVLAPEYSAKSGAAHL 926

949 myfsyagcfrfaylvahklsfedvllvfsavfgamavgygsfapayakksaahl 1008

927 FALLEKKPNIDRSQGGKRPDCGNLERRREVSEFRCRDPVFLRGLSLISIERGTVAF 986

1009 lmlektplldyslsteqlmnpntlegvntfgevfnyptrcpdlvplqslslevkkgqllal 1068

987 VGSAGGKSTVOLQRLQRLDPVOGQVLFDGVDKALENVMQLSQAIVQOEPVLFMCSTA 1046

1069 vassgsgkstvtvqlerfcdpdiagkvlldgkeltkrlnvgvlrahgylvsqepdlffcsia 1128

1047 ENIAYGDNRSVVPPLDEIKERANANIHSEGLEPERYNTVOGLKGAQLSGGOKRLAIAR 1106

1129 enlaygdnsrvvsgeelivraakeanahafeslpxkystkvdkgqlsggqqrataral 1188

1107 ALLQPKLILDEATSLVNDSEKVVQVHALDKARTRGRLTVYTHRSALQONMDLYVLEIN 1166

1189 alvrpghlilldeatsaldeseavvgealdakaregtrctivlahrlstlqnadllvfvqn 1248

1167 GKIKEGTQHQLLRNDRDIYFKLVNAQS 1193

1167 GKIKEGTQHQLLRNDRDIYFKLVNAQS 1193



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:32:48 ; Search time 42.7 Seconds  
(without alignments)  
2131.817 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012

Sequence: 1 MIIIGILASLVNGACLPIMPL.....QELLRRNDYFKLVNQSVQ 1195

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3573	59.4	1276	1	DVHYIC
2	3554	59.1	1280	1	DVHUI
3	3526.5	58.7	1276	1	DVMSI
4	3520.5	58.6	1276	1	A34786
5	3504.5	58.3	1279	1	DVHUI
6	3480.5	57.9	1276	1	DVMS2
7	3476.5	57.8	1277	2	JH0502
8	3456	57.5	1287	2	SS5692
9	3441.5	57.2	1281	2	I48123
10	3428.5	57.0	1278	2	S41646
11	3264	54.3	1104	1	DVMSIA
12	2839	47.2	1321	2	T42842
13	2835	47.2	1321	2	T42228
14	2529	42.1	1294	2	T19982
15	2380.5	39.6	1275	2	T31073
16	2351.5	39.1	1286	2	T02187
17	2339	38.9	1292	2	T48007
18	2328	38.7	1321	2	T23476
19	2326.5	38.7	1321	2	A47377
20	2326	38.7	1321	2	S27337
21	2315	38.5	1278	2	EB6155
22	2308.5	38.4	1302	2	A41249
23	2306	38.4	1230	2	EB5023
24	2290	38.1	1229	2	DB5023
25	2287	38.0	1229	2	T52319
26	2269	37.7	1229	2	F66155
27	2229.5	37.1	1323	2	H85202
28	2219.5	36.9	1266	2	T22090
29	2215	36.8	1302	2	B41249

30	2190	36.4	1310	2	S30328	multidrug resistan
31	2184.5	36.3	1268	2	T22094	hypothetical prote
32	2166	36.0	1307	2	T30882	multidrug resistan
33	2153	35.8	1286	2	A42150	P-glycoprotein pgp
34	2147	35.7	1245	2	G66404	probable P-glycopr
35	2135.5	35.5	1254	2	S27338	P-glycoprotein C -
36	2127.5	35.4	1222	2	T14805	hypothetical prote
37	2119.5	35.3	1247	2	F66405	probable P-glycopr
38	2119	35.2	1302	2	S30327	multidrug resistan
39	2116	35.2	1408	2	T43261	multidrug resistan
40	2107.5	35.1	1362	2	T41534	leptomycin B resis
41	2063	34.3	1233	2	T04251	P-glycoprotein 2 -
42	2054	34.2	1254	2	T30855	multidrug resistan
43	2022	33.6	1318	2	T21266	hypothetical prote
44	2014.5	33.5	1327	2	T21268	hypothetical prote
45	2004.5	33.3	1283	2	T18938	hypothetical prote

## ALIGNMENTS

RESULT 1  
DHYIC  
multidrug resistance protein 1 - Chinese hamster  
N:Alternate names: P-glycoprotein pgp1  
C:Species: Cricetus griseus (Chinese hamster)  
C>Date: 31-Dec-1990 #sequence\_revision 30-Sep-1992 #text\_change 19-Jan-2001  
C:Accession: A38696; C38696; B38696; A27126; S33768; I32823  
R:Devine, S.E.; Hussain, A.; Davide, J.P.; Melera, P.W.  
J. Biol. Chem. 266, 4545-4555, 1991  
A>Title: Full length and alternatively spliced pgp-1 transcripts in multidrug-resista  
A:Reference number: A38696; MUID:91154265  
A:Accession: A38696  
A:Molecule type: mRNA  
A:Residues: 1-1276 <DEV>  
A:Cross-references: GB:M59253; NID:g191154; PIDN:AAA37004.1; PID:g191155  
A:Accession: C38696  
A:Molecule type: mRNA  
A:Residues: 108-1276 <DE1>  
A:Cross-references: GB:M59254; NID:g191156; PIDN:AAA37005.1; PID:g191157  
A:Experimental source: clone ADX185  
A:Accession: B38696  
A:Molecule type: mRNA  
A:Residues: 1-32,771-1276 <DE2>  
A:Cross-references: GB:M59252; NID:g191152; PIDN:AAA37003.1; PID:g191153  
A:Experimental source: clone ADX124  
R:Endicott, J.A.; Juranka, P.F.; Sarangl, F.; Gerlach, J.H.; Deuchars, K.L.; Ling, V.  
Mol. Cell. Biol. 7, 4075-4081, 1987  
A>Title: Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chines  
A:Reference number: A27126; MUID:88122132  
A:Accession: A27126  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 706-1276 <END>  
A:Cross-references: GB:M17897; NID:g191158; PIDN:AAA37006.1; PID:g191159  
R:Zastawny, R.L.; Ling, V.  
Biochim. Biophys. Acta 1173, 303-313, 1993  
A>Title: Structural and functional analysis of 5' flanking and intron 1 sequences of  
A:Reference number: S33768; MUID:93305724  
A:Accession: S33768  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-21 <ZAS>  
A:Cross-references: EMBL:L03286  
R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.  
Cell Growth Differ. 2, 429-437, 1991  
A>Title: Analysis of the Chinese hamster P-glycoprotein/multidrug resistance gene pgp  
A:Accession: I52823  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-21 <RES>  
A:Cross-references: GB:S81975; NID:g240862

C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell  
structurally and functionally unrelated lipophilic antitumor drugs.

C:Genetics:

A:Gene: ppg1

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding; P-

F:407-601/Domain: ATP-binding cassette homology <ABC1>

F:424-431/Region: nucleotide-binding motif A (P-loop)

F:548-552/Region: nucleotide-binding motif B

F:1050-1246/Domain: ATP-binding cassette homology <ABC2>

F:1067-1074/Region: nucleotide-binding motif A (P-loop)

F:1193-1197/Region: nucleotide-binding motif B

F:87,91,96/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:430/Binding site: ATP (Lys) #status predicted

F:1073/Binding site: ATP (Lys) #status predicted

Query Match 59.4%; Score 3573; DB 1; Length 1276;

Best Local Similarity 55.2%; Pred. No. 7.6e-213;

Matches 679; Conservative 248; Mismatches 259; Indels 44; Gaps 8;

```

QY 1 MIIIGLILVNGALPLPLVGLGMSDNLISGLVOTN-----
DB 50 MLVGTLLAIIHGVALPLMMVLFEGDMDTDFASVGNIPFNATNATQVNASDIFGLKEEMT 109
QY 39 TYSEFRLTYVYGIGVALIFGYIOISLMTTARQTRIKRPFHSYLAODIGMFDSCD 98
DB 110 TYAA-----YTTGIGAVLVIAIYQVSWFCLAGROIHKIKQFHAIIMNDEIGMFDVHD 164
QY 99 IGEINTMT-DIDKISDQIGKIALFQNNSTFSGIAGLVGKMKLTVLTSTSPILMA 157
DB 165 VGEINTRLTLDVSKINQIGKMFQAMATPFGGFIPTKMGKLTLLVLAISVGL 224
QY 158 SAACSRMIVLSLTKSELAYSAGAAVEVLSIRTYIAEPQEKELORYONIKDAKF 217
DB 225 SAGIMAKILSFTTKEIOAYAKAGAAVEVLAIRTYIAEGQKKEERINNNLEAKRL 284
QY 218 GIKRTIASKVSIGLAVYFEMNGTGYIAFWGTSLLNGEPYTGITVLAFFSVHSYCI 277
DB 285 GIKKAITANISMGAFLIIVASVALAFWYGSIVLSKE--YSIQGLVTFEFAVLIAFESI 342
QY 278 GAAPVHETPAIANGAFHITQVLDKRPISIDNESTAGYKPSIGTYEFKAVSNYSRP 337
DB 343 GQASPNTIEAFANAGAAVEFIINIDNKRPSIDFSKNGKPNINGNLEFKNIHSPYSRK 402
QY 338 SIKLIKGLNLRKSGEVALVGLNGSGKSYVOLLRXYDDDEFIWDEMDIALNRRH 397
DB 403 DVQILKGLNLKVGSGQVVALVGNCGSGSTVQLRLYDFTEGVSTIDGODITINRY 462
QY 398 YRDHIGVAVSOEPLFGTTISNNIKYGRDVTDEMERARERANAYDFIMEPNKFNELVG 457
DB 463 LREIIGVAVSOEPLFATTIAINIRYGRNVTMDIEKAVKANAYDFIMKPHFEDTLVG 522
QY 458 EKGQMSGGQKORIAIARALVNRPKIILDEATSAIDSESKSAVOALEKASKRTIIV 517
DB 523 ERGQISGGQKORIAIARALVNRPKIILDEATSAIDSESAVVOALDKAREGRTIIV 582
QY 518 AHRISTRSADLIYTLKDGMAEKGAAHLEMAKRGYLSVMSQ-----DIKKAD 567
DB 583 AHRISTRNADIIYLFQGGVIVEQGNHEELREKGIYFKLVTQTAGNEIELGNEVSGK 642
QY 568 EOMESMTYSTERKTNSTL-PLHSYKSIKSDFIKAEEST-OSKEISLPEVSLIKLTKNK 624
DB 643 NEIDNLMSSKSDASSLRRSTRSRIRGPHDQRLSTKALDEDDVPPIFWMLIKNS 702
QY 625 PEMFPVVLGTSLAVLNGTVHPVPSIIFAKITMP-GNNDKTKLKHDAETYSMIFVILGI 683
DB 703 SEMYFVVGIFCAIVNGALQPAFSLIISKVYGVFTRNDDDETKHSLFLLILGLVI 762
QY 684 CFVSYFMQGLFYRAGELTMRHLHIAFKAMLYODIAMDEKENSSTGLTTLIADIAOI 743
DB 763 SFTTFYLOGFTFGAGELTIRKLRYMFKSKMLKODVSWDNPKVTGALTLLRLANDAGOV 822

```

```

QY 744 QGATGSRIGVLTQATNATMGSLVSIISFYGMEFTLILSTAPVLAVTGMEFTAAVTGANK 803
DB 823 KGATGRLAVLTQATNATMGSLVSIISFYGMEFTLILSTAPVLAVTGMEFTAAVTGANK 882
QY 804 DKQRLKAGTATPALTENITIVSLTREKAPFQMYEEMLOTOHNNKSKNOIIGSCAFS 863
DB 883 DKRLKESGKATPALTENITIVSLTREKAPFQMYEEMLOTOHNNKSKNOIIGSCAFS 942
QY 864 HAFYFAYAGFRGAVLIQAGRTPEGMFTVFAIYAGMAIGKTVLAPEYKASKA 923
DB 943 QAMMYESTACFRGAVLIQAGRTPEGMFTVFAIYAGMAIGKTVLAPEYKASKA 1002
QY 924 AHLPALLEKPNIDSRQSEKPKDPCGNLEFREVSYPCRPDVLTLRLSLSTERKT 983
DB 1003 SHITMIETKPSIDSYSTGKLKPTLEGNVKFNENVENYFPRDIPVQLGINTLEVKKGQT 1062
QY 984 VAFYSSGCGKSTVQLRLYDPVQGVLEFDGDAELVQMLRSQIALVPEPVLFC 1043
DB 1063 LALVSSGCGKSTVQLRLYDPVQGVLEFDGDAELVQMLRSQIALVPEPVLFC 1122
QY 1044 SIENIAYGDSRVVPRDEIKRANANHSFIEGLPEKTYOVLGKGAQSGQKORLA 1103
DB 1123 SIENIAYGDSRVVPRDEIKRANANHSFIEGLPEKTYOVLGKGAQSGQKORLA 1182
QY 1104 IARALQPKLILDEATSAIDNESEKVOHALDKARTGTCVTRHLSAIONADLIIV 1163
DB 1183 IARALVQPHILILDEATSAIDNESEKVOHALDKARTGTCVTRHLSAIONADLIIV 1242
QY 1164 LHNKRIKQEGHQLNRRIDYFKLVNAOS 1193
DB 1243 IONGKVRKHEGTHOOLAKQIFYSMSVOA 1272

```

## RESULT 2

```

Multidrug resistance protein 1 - human
N:Alternate names: P-glycoprotein 1
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence, revision 18-Aug-1995 #text, change 19-Jan-2001
C:Accession: A34914; PS0162; S15500; A25059; SA3838; I52238; I65204
R:Chen, C.; Clark, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
J. Biol. Chem. 265, 506-514, 1990
A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
A:Reference number: A34914; MUID:90094448
A:Accession: A34914
A:Molecule type: DNA
A:Residues: 1-1280 <CH>
A:Cross-references: GB:M29447; GB:J05168; NID:9187496; PIDN:AAA59576.1; PID:9386862
R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to JIPID, April 1991
A:Reference number: PS0162
A:Accession: PS0162
A:Molecule type: DNA
A:Residues: 1-22 <KID>
R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to the EMBL Data Library, April 1991
A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) express
A:Reference number: S15500
A:Accession: S15500
A:Molecule type: DNA
A:Residues: 1-22, 'R' <K12>
A:Cross-references: EMBL:X58723; NID:934522; PIDN:CAA41558.1; PID:934523
R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson,
Cell 47, 381-389, 1986
A:Title: Internal duplication and homology with bacterial transport proteins in the m
A:Reference number: A25059; MUID:87028230
A:Accession: A25059
A:Molecule type: mRNA
A:Residues: 1-184, 'V', 186-1280 <CH2>
A:Cross-references: GB:M14758; NID:9187468; PIDN:AAA59575.1; PID:9307180
R:Chambers, T.C.; Pohl, J.; Glass, D.B.; Kuo, J.F.
Biochem. J. 299, 309-315, 1994
A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase

```

A:Reference number: S43838; MUID:94220047  
A:Accession: S43838  
A:Molecule type: protein  
A:Residues: 656-689 <CHA>  
R:Geisler, V.; Weger, S.; Probst, H.  
Biochem. Biophys. Res. Commun. 169, 796-802, 1990  
A:Title: md1/p-glycoprotein gene segments analyzed from various human leukemic cell lit  
A:Reference number: 152238; MUID:9029529  
A:Accession: 152238  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 178-215 <RES>  
A:Cross-references: GB:M37724; PIDN:AAA8047.1; PID:G553314  
A:Accession: 165204  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 800-856 <RE2>  
A:Cross-references: GB:M37725; NID:G183538; PIDN:AAA8048.1; PID:G553315  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell  
structurally and functionally unrelated lipophilic antitumor drugs.  
C:Genetics:  
A:Gene: GDB:PGY1, MDR1  
A:Cross-references: GDB:120712; OMIM:171050  
A:Map position: 7q21-7q21  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein;  
F:1-638, 653-1280/Region: duplication  
F:49-350/Domain: hydrophobic <HBL>  
F:351-637/Domain: hydrophilic <HL1>  
F:410-604/Domain: ATP-binding cassette homology <ABC1>  
F:427-644/Region: nucleotide-binding motif A (P-loop)  
F:551-555/Region: nucleotide-binding motif B  
F:638-708/Domain: linker <LIN>  
F:709-993/Domain: hydrophobic <HB2>  
F:994-1280/Domain: hydrophilic <HL2>  
F:1053-1249/Domain: ATP-binding cassette homology <ABC2>  
F:1070-1077/Region: nucleotide-binding motif A (P-loop)  
F:1196-1200/Region: nucleotide-binding motif B  
F:91.94.99/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:433/Binding site: ATP (Lys) #status predicted  
F:667, 667, 671/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #statu  
F:1076/Binding site: ATP (Lys) #status predicted

Query Match 59.1%; Score 3554; DB 1; Length 1280;  
Best Local Similarity 55.6%; Pred. No. 1.1e-211;  
Matches 682; Conservative 246; Mismatches 263; Indels 36; Gaps 9;

1 MTLGLASLVNGACLPIMPLVVGEMSD-----NLTSGLYOF--NTYSF----- 43  
51 MVTGLTAAITIHAGLPMLLVGEMTDIPFANGNLEDLNTNSNDINDTGFNNLEED 110  
44 --RLTVYVGIVAAITFEGYIOISLITTAAROTKRIRKOPFHSVLAODIGFNSCDICE 101  
11 MRRVAYVYSGIAGVAVAYIQVSWCLAGRIKIRKOPFHAIKROELIGFNDVADVE 170  
102 LMTFRMT-DIDKISDGIKIALIFONMSTFSGIAGLVKCKKLLVLTSTPLMASAA 160  
171 LMTRLTLDVSKINEGIDKIGMFQSMATFETGFTVGRGKLLVLAISPVLGSLAA 230  
161 ACSRWVISTLSKEASVAGAAVEEVLISRTVIAFRKOEKLOLYTONLMDAKDFGIK 220  
231 VAKLITSTDTKELLAYAKAGAAVEEVLIAKRTVIAFGCKKELEYKNKLEBAKRTIGIK 290  
221 RTIASKVSLGAVYFPMNGTYGLAFMYGTSLLNGEPGYTIGTVLAFVFSVHSSYCIGA 280  
291 KATIANISIGAFLLIYASVAFMYGTTLVSGE--YSIGVLYVFFSVLLIGAFSVQGA 348  
281 VHFPEFALARAQAHIPOVIDKKSINDFNSTAGIKPESIECTVEKKNVSNPSPRSK 340  
349 SPSEIAFANARAAAEIKFIIDNKPSIDSYSGKHKPDNIKGLERFNVHFSYPSRKEVK 408

341 ILKGLNLRKSGEYVALVGLANGSGKSTVQLQRLYPDDEGIMVNDENDIRALNVRYRD 400  
409 ILKGLNLRKSGEYVALVGLANGSGKSTVQLQRLYPDDEGIMVNDENDIRALNVRYRD 468  
401 HIGVSOEPLVEFGTTISNNIKYGRDVTDEEMERAAANAVDTIMEPNKNTLVGEKG 460  
469 IIGVSOEPLVEFGTTISNNIKYGRDVTDEEMERAAANAVDTIMEPNKNTLVGEKG 528  
461 AQMSGOKORIAIARALVNPVKILLDEATSLDSEKSAVOALAEKASGRTTIVVHR 520  
529 AOLSGOKORIAIARALVNPVKILLDEATSLDSEKSAVOALAEKASGRTTIVVHR 588  
521 ISTISADITVILKGMALAEKHAELAKRGLYSLVMSQ-----DIKKADEOMESMT 574  
589 ISTVKNADYIAFDDGVYVEKGNHDELMKEKTIYKRLVYMQNAGNVEELNADSEKSCI 648  
575 YSTERKTNLSPLSHV-----SKIR-SDFIDKAEBSSTOSKEISLDEVSILKTLKRPW 627  
649 DALEMSNDSRSSILTRKSTRSVKRSQADKRLSTKEALDESIPPVSFWIRIMKLNLEW 708  
628 PRVVLGTLASVANGVHPVFSTIFPAKITMPGN-NDKTLTKDAEYISMIFVLGICPV 686  
709 PRVVLGTLASVANGVHPVFSTIFPAKITMPGN-NDKTLTKDAEYISMIFVLGICPV 768  
687 SYFMGLFYGRAGELITMRLRLAFKAMLYODIAWFDEKENSTGGLTTLTADIAIOGA 746  
769 TFFLOGFTFGKGEITITKRLRYMFRSMRLRODVSNFDDKNTGALTTFLLANDAAQVGA 828  
747 TGSRLGVLTONATNGLSVIISFTYGMETFLILSTAPLAVYGMETFAAMTGFANKDQ 806  
829 ISSRLAVIQNIAKNGTGLTISFTYGMETFLILSTAPLAVYGMETFAAMTGFANKDQ 888  
807 ELKHAQKATLEALENRTIVSLTRKAPFQMYEMQOHRMTSKAQTIGSCYAFSAF 866  
889 ELKHAQKATLEALENRTIVSLTRKAPFQMYEMQOHRMTSKAQTIGSCYAFSAF 948  
867 IYFAAAGFRFAYLIQAGMTPEGMFYFTAIAYGAMAIGTLYLAPESKAGGAHL 926  
949 MFTSAGCRFRFAYLIQAGMTPEGMFYFTAIAYGAMAIGTLYLAPESKAGGAHL 1008  
927 FALLEKRPIDRSOEGKKPDTCEGNLEFREVSYFYPCRPDVFILGSLSTIERKTVAF 986  
1009 IMIERTPLIDSYSTEGILMPNLTLEGVTFGEVFNPRPDIPVLQGLSLLEVKKGOTLAL 1068  
987 VSSSGCKSTVQLQRLYPDDEGIMVNDENDIRALNVRYRD 1046  
1069 VSSSGCKSTVQLQRLYPDDEGIMVNDENDIRALNVRYRD 1128  
1047 ENIAYGDNRRVYPLDEIKFAAANAHSFIEGPEKYNQVGLKAGOLSGGOKORLAIAR 1106  
1129 ENIAYGDNRRVYPLDEIKFAAANAHSFIEGPEKYNQVGLKAGOLSGGOKORLAIAR 1188  
1107 ALLQRPKILLDEATSLDSEKSAVOALAEKASGRTTIVVHR 1166  
1189 ALVROPKHILLDEATSLDSEKSAVOALAEKASGRTTIVVHR 1248  
1167 GKIKDGHQHELLRNNDIYFKLVNMS 1193  
1249 GRVKEHGHQHELLRNNDIYFKLVNMS 1275

RESULT 3  
DYMST  
multidrug resistance protein 1 - mouse  
N:Alternate names: P-glycoprotein 1  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text\_change 19-Jan-2001  
C:Accession: A33719; A25057; I57510  
R:Raymond, M.; Gros, P.; U.S.A. 86, 6488-6492, 1989  
A:Title: Mammalian multidrug-resistance gene: correlation of exon organization with s  
A:Reference number: A33719; MUID:89367274  
A:Accession: A33719

A:Molecule type: DNA  
 A:Residues: 1-1276 <RAY>  
 R:Gros, P.; Croop, J.; Hausman, D.  
 Cell 47, 371-380, 1986  
 A:Title: Mammalian multidrug resistance gene: complete cDNA sequence indicates strong homology to the Drosophila multidrug resistance gene  
 A:Reference number: A25057; MUID:87028228  
 A:Accession: A25057  
 A:Molecule type: mRNA  
 A:Residues: 1-1276 <GRO>  
 A:Cross-references: GB:M14757; NID:g199100; PIDN:AAA9005.1; PID:g387426  
 R:Raymond, M.; Gros, P.  
 Mol. Cell. Biol. 10, 6036-6040, 1990  
 A:Title: Cell-specific activity of cis-acting regulatory elements in the promoter of the multidrug resistance gene  
 A:Reference number: I57510; MUID:91042535  
 A:Accession: I57510  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-21 <RES>  
 A:Cross-references: GB:M60348; NID:g199102; PIDN:AAA39513.1; PID:g554199  
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell lines  
 C:Genetics:  
 A:Gene: mdr1 (pgp1)  
 A:Insertions: 21/2; 38/3; 92/1; 111/2; 175/2; 233/3; 274/2; 332/3; 370/3; 449/3; 517/3  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane protein  
 F:1-637/653-1276/Region: duplication  
 F:409-603/Domain: ATP-binding cassette homology <ABC1>  
 F:426-433/Region: nucleotide-binding motif A (P-loop)  
 F:530-554/Region: nucleotide-binding motif B  
 F:1051-1247/Domain: ATP-binding cassette homology <ABC2>  
 F:1068-1075/Region: nucleotide-binding motif A (P-loop)  
 F:1194-1198/Region: nucleotide-binding motif B  
 F:72,91,96,103/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:432/Binding site: ATP (Lys) #status predicted  
 F:1074/Binding site: ATP (Lys) #status predicted

Query Match 58.7%; Score 3526.5; DB 1; Length 1276;  
 Best Local Similarity 55.1%; Pred. No. 5.7e-210;  
 Matches 678; Conservative 24; Mismatches 262; Indels 49; Gaps 10;

QY 1 MIIIGTASLVNAGCLPMLPLVGLGMSDNL-----SGLVQNT----- 39  
 DB 50 MIIIGTASLVNAGCLPMLPLVGLGMSDNL-----SGLVQNT----- 39  
 QY 40 ---SFRRLTYVYVIGVAAALIFGVIQISLWITAAARQTRKRRQFPVSVALAQDIGNFDS 96  
 DB 110 MALTAY-----YTGIGAGVLTVAIVDSIMCLAGROIHKIRKFFHAINNOEIGMFDV 164  
 QY 97 CDIGELNTRMT-DIDKISDGIKRIALFQNMSTFSGIAGLVKGMKLTLTSTPLI 155  
 DB 165 HDVGLNTRMT-DIDKISDGIKRIALFQNMSTFSGIAGLVKGMKLTLTSTPLI 155  
 QY 216 DFGIKRTIASVSGAVYFFMNGTYGLAFWGTSLINGEGCYITIGYLAFFSVIHSY 275  
 DB 285 NVGIKRTIASVSGAVYFFMNGTYGLAFWGTSLINGEGCYITIGYLAFFSVIHSY 275  
 QY 276 CIGAAVPHFEFFAARGAFFHFOVNDKPPSIDNFTAGYKPEISIEGVEKFNVSFNP 335  
 DB 343 SIGHAPRIEAFANARGAFFHFOVNDKPPSIDNFTAGYKPEISIEGVEKFNVSFNP 335  
 QY 403 RSEVQITLGLMLKVKSGQVALVNSGCKSTVQMLQRLDPLEGVVSDIGQDIRTNV 462  
 QY 396 RHYVDHIGVVSQEPVLTSTISNNIKGRDVTDEEMKARARNAADFMEFNKFTL 455  
 DB 463 RYLRREIIGVVSQEPVLTSTISNNIKGRDVTDEEMKARARNAADFMEFNKFTL 455

QY 456 VGEKAGMSGQKORAIARALVNPRIILLDEATSAIDSEKSAVAOLEKASGRFTI 515  
 DB 523 VGEKAGMSGQKORAIARALVNPRIILLDEATSAIDSEKSAVAOLEKASGRFTI 515  
 QY 516 VVAHRLSTIASADILVTLKGMALAEKGAHLEAKKRLVSLVKSQDIKADQEMMTY 575  
 DB 583 VIAHRLSTIASADILVTLKGMALAEKGAHLEAKKRLVSLVKSQDIKADQEMMTY 575  
 QY 576 STEKRTSLPLHSVSKSPIDKA-----EESTQKESL-----PEVSLKILK 621  
 DB 642 GSQSDTDASELTSEES-KSLPIRSIYSVRHKDDQRRLSMKAEVDEVPVLSFRIIN 700  
 QY 622 LNKEMPEVVLGTASVLTGVPVHPFSIIPAKITMG-NNDKTTLKHADEIYSMTFVL 680  
 DB 701 LNSEMPYLLVGLCANVINCIOVFAIVPSRIYGVSRDDDEHETKQNCNLSFLPLVM 760  
 QY 681 GVICFVSFEMQGLFYGAGELTMRHLAFKAMLYODIAMFEDEKSTGGLTILADI 740  
 DB 761 GLISFVTFEGGFEGKAGEILTFRVRYMYFKSMLRODISWFDHKNKSTGLTTRLSA 820  
 QY 741 AQIOGATGSRIGVLTQATNMGSLVSIISFYGEMTEFLIISIAVPLAVTGMIEAATGF 800  
 DB 821 SVYAGAGARLAVTQVANGLGVILSLVGMQTLTLLVILPLVGLGIEKLSGQ 880  
 QY 801 ANKDQELKHAGIATPALENIPTIVSLTREKAEQOMEYEMLOTHRTSKKAOIGSCY 860  
 DB 881 ALNKKQLEISGKATATPALENIPTIVSLTREKAEQOMEYEMLOTHRTSKKAOIGSCY 860  
 QY 861 AFSHAIFYFAVAGFRGAVLIQAGMTPEGMETVFAAIYAGAMAIKTLVLAPYSKAK 920  
 DB 941 SFTQAMMYFSAACFRGAVLIQAGMTPEGMETVFAAIYAGAMAIKTLVLAPYSKAK 920  
 QY 921 SGAAHLLALLEKKNIDRSQEGKKPTCEGNEFRVSFFPCRPVFLIRGLSLIER 980  
 DB 1001 VSASHILIRITKEPEDIYSSTEGKPTLLEGNKFNQVQNPFRPPIVQGLSLVYK 1060  
 QY 981 GKTVAFVSSGCKSTSVOLRLDLYDPVQGVLEFDGVAKEALNOMLRSAIYPOBPVL 1040  
 DB 1061 GQTLALVSSGCKSTSVOLRLDLYDPVQGVLEFDGVAKEALNOMLRSAIYPOBPVL 1040  
 QY 1041 FNCISAEINAYGDSRVVPLDEIKEAANANHSFIEGLDEKTYOGLGAQSGSQKQ 1100  
 DB 1121 FDCISAEINAYGDSRVVPLDEIKEAANANHSFIEGLDEKTYOGLGAQSGSQKQ 1100  
 QY 1101 RLATARALVROPHTLLDEATSAIDSEKSAVAOLEKASGRFTIYIARHLSIQADL 1160  
 DB 1181 RLATARALVROPHTLLDEATSAIDSEKSAVAOLEKASGRFTIYIARHLSIQADL 1160  
 QY 1161 IVVLNKGIRKQGTQOELLRNRDIYFKLVNA 1191  
 DB 1241 IVVLNKGIRKQGTQOELLRNRDIYFKLVNA 1191

RESULT 4  
 A34786  
 multidrug resistance protein 1a - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 02-Feb-2001  
 C:Accession: A34786; A35671  
 R:Devault, A.; Gros, P.  
 Mol. Cell. Biol. 10, 1652-1663, 1990  
 A:Title: Two members of the mouse mdr gene family confer multidrug resistance with ov  
 A:Reference number: A34786; MUID:90205845  
 A:Accession: A34786  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1276 <DEV>  
 A:Cross-references: GB:M30697; NID:g199111; PIDN:AAA39517.1; PID:g387429  
 R:Hau, S.I.H.; Cohen, D.; Kirschner, L.S.; Lotenstein, L.; Hartstein, M.; Horwitz, S.B.  
 Mol. Cell. Biol. 10, 3596-3606, 1990  
 A:Title: Structural analysis of the mouse mdr1a (P-glycoprotein) promoter reveals the



A:Accession: A335871  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-525, 'HV', 528-938, 'S', 940-1035, 'V', 1037-1276 <HSD>  
 A:Cross-references: GB:M3581; NID:g199104; PIDN:AA39514.1; PID:g387427  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-loop  
 F:406-600/Domain: ATP-binding cassette homology <ABC1>  
 F:423-431/Region: nucleotide-binding motif A (P-loop)  
 F:547-551/Region: nucleotide-binding motif B  
 F:1049-1245/Domain: ATP-binding cassette homology <ABC2>  
 F:1066-1074/Region: nucleotide-binding motif A (P-loop)  
 F:1193-1196/Region: nucleotide-binding motif B  
 F:429/Binding site: ATP (Lys) #status predicted  
 F:1072/Binding site: ATP (Lys) #status predicted

Query Match 58.6%; Score 3520.5; DB 2; Length 1276;  
 Best Local Similarity 54.3%; Pred. No. 1.3e-209;  
 Matches 668; Conservative 251; Mismatches 265; Indels 47; Gaps 8;

QY 1 MIIIGTILSYNGACPLPLMLVLEMSDNLSGCLVOTN-----T 39  
 DB 50 MLVGTIAIITHGVALLPLMLIFGDMTDSFASVGNVSKNSTNMEADKRAMFAKLEEMTT 109  
 QY 40 YSFRLTLVYGVIGVALLFGVYQISLMTITTAARQTKTRKOPFHSVLADIDMPSCDI 99  
 DB 110 YAV-----YTTIGAGVLLVAVYQVSWCLAAAGROTHKROKFFHAIMDGEIMFVHDV 164  
 QY 100 GELNFRMT-DIDKISDIDGDKIALLFQNMSTFSGVLAVGLVGMKLTLYLSTPLMAS 158  
 DB 165 GELNFRMTDVSQKINGIDGDKIMFQAMATFPGFIIIGTRGMKLTLYLSTPLMAS 224  
 QY 159 AAACSMVYISLTKSELAYSAGAAVEVLSIRTVIAFRQEKELQRTQNLKADKDFG 218  
 DB 225 AGIMAKILTSFTDKELHAAVAKAGAAVEVLAIRTVIAFGQKKELEERYNNNEAKRLG 284  
 QY 219 IKRTISKVSYGVYFMNGTVGLAFWYGSLLNCEPQTITVLAFFSVYHSYCTG 278  
 DB 285 IKKATLANISMGAFLLITASTALAWYGSLSVSK--YSIQVLTFFSVYIGAFSVG 342  
 QY 279 AAVPHETALARGAAPHIFOVYIDKRPISDNFSTAGYKPEISIGYVEFKVSNFYSRPS 338  
 DB 343 QASPNIEATFANARGAYEVFKIIDNKPISDSFSKCHKNIDGNELEFKNIHYSYRKE 402  
 QY 339 IKTLKGLNLRISGEFVALVGLNGSKSTYVQLLRYPDDGFIWDENDIALNVRHY 398  
 DB 403 VOILKGLNLRKVGQTVAVLVGNGSKSTYVQLMRLYPDLDMVSDIGODIRTINRYL 462  
 QY 399 RHDIQVSGSEPLVFGTTISNIIKYGRDQVTDDEMERARANAAYDTIMEPKNFNTLYGE 458  
 DB 463 RELTIGVSGSEPLVFGTTISNIIKYGRDQVTDDEMERARANAAYDTIMEPKNFNTLYGE 522  
 QY 459 KGAMSGGOKORIAIARALVLRNPKIILDEATSAIDSESKSAVQALAKSKARTTYVA 518  
 DB 523 RGAQLSGGOKORIAIARALVLRNPKIILDEATSAIDSESKSAVQALAKSKARTTYVA 582  
 QY 519 HRLSTIRSDLVTLTKDGLMAEKGAHAELMAKGLYSLVMSQ-----DIKKA 568  
 DB 583 HRLSTIRSDLVTLTKDGLMAEKGAHAELMAKGLYSLVMSQ-----DIKKA 642  
 QY 569 QMSMTYSTERKTNPLPLSVKSIKSDFT-----DKAESTQSKETSLPVSILKILKLN 623  
 DB 643 EIDNLMSSSDSGSSLIR--RSTRKSGCPHQDRLKSLKREKALDDVDPASPFWRLIKLN 700  
 QY 624 KPEMPVNLCTLASVLTNGVHPVFSIIPAKIITMFGN-NDKTLTKDAETYSMIFVLIG 692  
 DB 701 STMTPLFVVGIRCAITINGIOPAFSVIFSKVGVFTNGEPETQORONSLFLFLILGI 760  
 QY 683 ICFVSYFMQGLFYGRAGELITMLRLHLAFKAMLYODIAMEDEKNSGTGLTTLTIAID 742  
 DB 761 ISPTFEFLGTFPGKAGEIITKRLRYMVFPSMLRDVSWEDDKNTGALTTLRLANDAAQ 820

QY 743 ICGATGSRIGVLQONATNMGLSVISIFYCGEMTFLILSTAPVLAVTGMIETAAMTGFAN 802  
 DB 821 VKGATGSRILAVIPQNTANLGTGIIISLYGMQTLTLLAVPITAIAGVEMKMLSGQAL 880  
 QY 803 KDKQELKHAQKITEALENIRITVLSLREKAFBQMT EEMLOTQHRNTSKKAOIIGSCYAF 862  
 DB 881 KDKQELKHAQKITEALENIRITVLSLREKAFBQMT EEMLOTQHRNTSKKAOIIGSCYAF 940  
 QY 863 SHAFTFAVAGRFAGAYLIOAGRMPDEGMFYFTALAYGAMALGKTVLAPEYSAKSG 922  
 DB 941 TOAMMFYSVACRFAGAYLIOAGRMPDEGMFYFTALAYGAMALGKTVLAPEYSAKSG 1000  
 QY 923 AAILFPLLEKPNIDRSQKRPDCEGNLEFREVSVFPACPDVFIILGLSLSTERG 982  
 DB 1001 ASHTITRIETPDIETSTYQGLKPNMEGNVQSGEVFNTPRSPVVLGSLLEPKKG 1060  
 QY 983 TVAVGSSGGKSTVQLQRLYDPVQGOVLPFGVDPAKELNQMRLSQATVBOEPVFN 1042  
 DB 1061 TLALVSSGGKSTVQLQRLYDPVQGOVLPFGVDPAKELNQMRLSQATVBOEPVFN 1120  
 QY 1043 CSTAENIAGDNRVYPLDEIKRANANAHSTIEGPEKYNQVGLKQSLGSGQKRL 1102  
 DB 1121 CSTAENIAGDNRVYPLDEIKRANANAHSTIEGPEKYNQVGLKQSLGSGQKRL 1180  
 QY 1103 AIALLOKPKIILLDEATSAIDSESKSAVQALAKSKARTTYVA 1162  
 DB 1181 AIALLOKPKIILLDEATSAIDSESKSAVQALAKSKARTTYVA 1240  
 QY 1163 VLHNGRIKEQTHQELLRNDRIFYEKLVAQS 1193  
 DB 1241 VLHNGRIKEQTHQELLRNDRIFYEKLVAQS 1271

RESULT 5  
 DIVERS  
 N:multidrug resistance protein 3 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1990 #sequence, revision 31-Dec-1990 #text, change 19-Jan-2001  
 C:Accession: J50051; S01346; A42213; I38015  
 R:van der Bielek, A.M.; Koolman, P.M.; Schneider, C.; Borst, P.  
 Gene 71, 401-411, 1988  
 A:Title: Sequence of mdm3 cDNA encoding a human P-glycoprotein.  
 A:Reference number: J50051; M01D:89138016  
 A:Accession: J50051  
 A:Molecule type: mRNA  
 A:Residues: 1-1279 <VAL>  
 A:Cross-references: GB:M23234; NID:g187501; PIDN:AA36207.1; PID:g307181  
 R:van der Bielek, A.M.; Baas, F.; ten Houte de Lange, T.; Koolman, P.M.; van der Velde  
 EMO J. 6, 3325-3331, 1987  
 A:Title: The human mdm3 gene encodes a novel P-glycoprotein homologue and gives rise  
 A:Reference number: S01346; M01D:8811519  
 A:Accession: S01346  
 A:Molecule type: mRNA  
 A:Residues: 856-1093, 'FVDFGFO', 1094-1279 <VA2>  
 A:Cross-references: EMBL:X06101; NID:g34524; PIDN:CAA29547.1; PID:g34525  
 R:Linke, C.R.; Smit, J.J.M.; van der Velde-Koerts, T.; Borst, P.  
 J. Biol. Chem. 266, 5303-5310, 1991  
 A:Title: Structure of the human MDR3 gene and physical mapping of the human MDR locus  
 A:Reference number: A42213; M01D:91161629  
 A:Accession: A42213  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-1279 <LIN>  
 R:Smit, J.J.; Mol, C.A.; van Deemter, L.; Wagenaar, E.; Schinkel, A.H.; Borst, P.  
 Biochim. Biophys. Acta 1261, 44-56, 1995  
 A:Title: Characterization of the promoter region of the human MDR3 P-glycoprotein gen  
 A:Reference number: I38015; M01D:95200972  
 A:Accession: I38015  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-61, 'RGSSRVDAQC', <RES>  
 A:Cross-references: EMBL:Z35284; NID:g1006662; PIDN:CAA84542.1; PID:g1006663

C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell  
 C:Genetics: structurally and functionally unrelated 11pophilic antitumor drugs.

A:Gene: GDB:PG3; MDR3

A:Cross-references: GDB:120713; OMIM:171060

A:Map position: 7q21-7q21

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr

F:1-640-653-1279/Region: duplication

F:412-606/Domain: ATP-binding cassette homology <ABC1>

F:429-436/Region: nucleotide-binding motif A (P-loop)

F:553-557/Region: nucleotide-binding motif B

F:638-694/Domain: linker <LINK>

F:1052-1248/Domain: ATP-binding cassette homology <ABC2>

F:1069-1076/Region: nucleotide-binding motif A (P-loop)

F:1195-1199/Region: nucleotide-binding motif B

F:91,97/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:435/Binding site: ATP (Lys) #status predicted

F:1075/Binding site: ATP (Lys) #status predicted

Query Match 58.3%; Score 3504.5; DB:1; Length 1279;  
 Best Local Similarity 55.2%; Pred. No. 1.3e-208;  
 Matches 676; Conservative 235; Mismatches 276; Indels 37; Gaps 8;

1 MIIIGLASLVNGACPLMPVLVGENSDNLSGCLVQNTYF---FRLT----- 47  
 57 MSLGIMIAIHSGSLPIMMIVGEMTKFVD---TAGFSPVNFSLSLNPKILERE 112  
 48 -----YVVGICVALLREYIOISLMTITANQTKRKIRKOPFHSVLAODIGFSCDICE 101  
 113 MTRVAVYVSGLAGVLAAYIOVSFWTLAAGKQIRKIRKOPFHALIROIGFMDINDTE 172  
 102 LMTFRMT-DIDKISDGGOKIALPQNMSPFSGVLAVGKGLTYLMTSPILMASA 160  
 173 LMTFRMTDIDKISDGGOKIALPQNMSPFSGVLAVGKGLTYLMTSPILMASA 232  
 161 ACSRWVLSLTKSELSKAGAVAEVLSISRTVIAFRAOEKELORYTONLKDARKFGIK 220  
 233 VAKILSAFSDKELAVAKAGAVAEALGAIKRTVIAFGQNKELRYQKHLNAEIGIK 292  
 221 RTIASKVSLGAVYFPMNGTYGLAFWKYGSLLINGEGYITIGVLAVFESVHSYCIGA 280  
 293 KAISANISMGIAFLILYASALAFWGSSTLYISKE--YTIAGNAVVFISLLIGASVGA 350  
 281 VHFPEFAIARGAHIFQVIDKKPSIDNEFSAKYPSEIECTVEKKNSEYPSRPSIK 340  
 351 APCIDAFANARGAAYVIFDIIDNPIDSEFSEGRHPDSIKGNLEFNDVHFSPSRANYK 410  
 341 ILKGLNLRTKSGEVALVGLNGSKSTVQLQRLYPDPDFIMVENDIRALNRYHND 400  
 411 ILKGLNLKQSGQYVALVSSGCKSTYQLQRLYPDPEGININDGQIRNFNNYLKE 470  
 401 HIGVVSQSEVLEFETISNNIKYGRDDYDEMEARAAREANAVDFIPEPNKNTLVGEGK 460  
 471 IIGVVSQSEVLEFETISNNIKYGRDDYDEMEARAAREANAVDFIPEPNKNTLVGEGK 530  
 461 AOMSGGOKORIAIARLVNPKILIDEATSADSEKSAVOALAEKASKGRTTVVAHR 520  
 531 AOMSGGOKORIAIARLVNPKILIDEATSADSEKSAVOALAEKASKGRTTVVAHR 590  
 521 LSTIRASADLVTLKDOMLAEKGAHAEKARGLYSLV-----MSODIKKAEQWE 571  
 591 LSTIRASADLVTLKDOMLAEKGAHAEKARGLYSLV-----MSODIKKAEQWE 650  
 572 SMYTSERKTNLPLHSVSKIK-SDFIDKA-BESTQSKESLSEVSLKILKLNPEMF 629  
 651 TMAPNAPNGSRLEFRSTQKLNKSNQKSLDVEDTGLEANPVPSFLKVLKNTKTEMPY 710  
 630 VVLGLTASVNLNGTVHVESIIFAKITMGNDKTLTKDAETIISNIFILVLCVSVAF 689  
 711 FVVGIVCAIANGLOPAFVSIFSEIIAIPGDDAVAKOOKINFIPLFLGLIISFETPF 770

690 MGLLEFGRAGELITWRLHLAFKAMLYODIAMFDEKENSTGGLTTLAIDAIQOGATGS 749  
 771 LQGFPGKAGELITWRLHLAFKAMLYODIAMFDEKENSTGGLTTLAIDAIQOGATGS 830  
 750 RIGVLTQATNKGSLVITISFYIEMFTLISIAVLAVALVGTMIETAAATGANKDKGELK 809  
 831 RLATIAQNIAMNGTGIISFYIEMFTLISIAVLAVALVGTMIETAAATGANKDKGELK 890  
 810 HAGKATATALENIRITVSLTDEKAEKMEYEMLOHNTGSKKQIIGSCYAFSAHYFE 869  
 891 AAGKATATALENIRITVSLTDEKAEKMEYEMLOHNTGSKKQIIGSCYAFSAHYFE 950  
 870 AYAAFGREGAVLLIAGMTPDEGMFTVETALVAGAMAGIKTVLAPEYSKASGAHLEAL 929  
 951 SVAGFERGAVLLIAGMTPDEGMFTVETALVAGAMAGIKTVLAPEYSKASGAHLEAL 1010  
 930 LEKKNIDRSQEGKKPOTCEGNEFEFVSPFPCRPVYFLIRGLISLSENGKYAFAFGS 989  
 1011 FERQPLIDYSSEGLKPKFEGNITFENVNYPRANVPVLAGSLSEVKKGQTLALGS 1070  
 990 SGGKSTSVQLQRLYPDVGQVLEFVDYDAKELANQWLRSGAIYPOBPVFNCSIAENI 1049  
 1071 SGGKSTSVQLQRLYPDVGQVLEFVDYDAKELANQWLRSGAIYPOBPVFNCSIAENI 1130  
 1050 AYGDNSRVPLDEIKKANAANINHSFIEGPEKXTYQVGLGAQLSGGOKRLAIARAL 1109  
 1131 AYGDNSRVPLDEIKKANAANINHSFIEGPEKXTYQVGLGAQLSGGOKRLAIARAL 1190  
 1110 QKPKILLDEATSAIDNDESEKVVQHALDKARTRGCLVYTRLSAIONADLVLAHNKI 1169  
 1191 RQPKILLDEATSAIDNDESEKVVQHALDKARTRGCLVYTRLSAIONADLVLAHNKI 1250  
 1170 KEGTHOELLNRDIFPKLVNAOS 1193  
 1251 KEGTHOELLNRDIFPKLVNAOS 1274

RESULT 6  
 DMS2  
 multidrug resistance protein 2 - mouse  
 N:Alternate names: P-glycoprotein MDR2  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
 C:Accession: A30409; S70711  
 M:Cell: P.; Raymond, M.; Bell, J.; Housman, D.  
 A:Title: Cloning and characterization of a second member of the mouse mdr gene family  
 A:Reference number: A30409; MUID:88302195  
 A:Accession: A30409  
 A:Molecule type: mRNA  
 A:Residues: 1-1276 <HSD>  
 A:Cross-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428  
 R:Kritschner, L.S.  
 Nucleic Acids Res. 24, 2829-2834, 1996  
 A:Title: De novo generation of simple sequence during gene amplification.  
 A:Reference number: S70711; MUID:96313253  
 A:Accession: S70711  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 43-92 <KIR>  
 A:Cross-references: EMBL:U46839; NID:g1228142; PIDN:AC52722.1; PID:g1228143  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
 C:Genetics: structurally and functionally unrelated 11pophilic antitumor drugs.

A:Gene: mdr2  
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane  
 F:1-637-653-1276/Region: duplication  
 F:409-603/Domain: ATP-binding cassette homology <ABC1>  
 F:426-433/Region: nucleotide-binding motif A (P-loop)  
 F:550-554/Region: nucleotide-binding motif B  
 F:1049-1245/Domain: ATP-binding cassette homology <ABC2>

Db 49 MALGTLAIHGTLLPLMLVEGYMTDSFTPSRDPHSDRAITNQSEINSTHTVSDTSLEE 108



Db 604 NANAAGFNGVYEGSGHKELMERGAVYFNLTQTVETSKDTEDEDETHIY--EKKLP 661  
QY 563 SLPLHS-----VKSISDPIDK-AEESTOSKEISLPEVSLKTLKANKPEMPV 630  
Db 662 VTHSTSNLVRKRSRNTJTSKVPETDEKVEDEEKKEGPPVPSPFKVMKLNKPEMPYF 721  
QY 631 VGTGLASVINGVHPVFSIIFAKIITMFGNNKTKLKHAEIYSMFEVLGVCIFPSYFM 690  
Db 722 VVGICAMINGATOPAFALIFSRITIGVFA-GPVSQMRSSSMYSLIFLALGVASFTEFL 780  
QY 691 QGLFYRAGEIITMRLRHLAFKAMLYODIAMDPEKENSGGLTTLIADIAIOGATGSR 750  
Db 781 QRFTEGKGEILTMRLRSGFSKSMRLQELGWFDDSKNSGALTTLRATDASVOGATGR 840  
QY 751 IGVLTQNTNMGLSVIISFYEMEMTFLIISAPVAVTGMETGAMTGPAKRDQELKH 810  
Db 841 LALLAQNVAANLGTALISFYGMQLTLTLAIAPVIAAGLVEMKMFAGHAKKKELEK 900  
QY 811 AGKITTEALENRTIYSLTRKAFQOMYEMLOTOHRNTSKAQIIGSCYAFSAFIYFA 870  
Db 901 AGKISTDAVLRITVAVSLTRERKFEAMERKSLGEPYRNSIKKAHLGLTYGLSQAHVLC 960  
QY 871 YAAGER-FGAYLIQAGRMTPCEGMFYFAIAYGAMAIGKTLVLAPEYSKAKSAHLFAL 929  
Db 961 ICMWFSVLGAVLYVSGMLKMLDEVLYSSAIVLGAMALGOTSSFADPYTKAMISAHIIFSL 1020  
QY 930 LEKKNIDRSQEGKKPDTCEGNLEPREVSPYPCRPDVFIILRGLSLSTERKTVAFVGS 989  
Db 1021 LERVQIDISYSDQEGKPKCSNVVFKGVNFVPTRPDITVLQGLDISYKQETLALVGS 1080  
QY 990 SCCGKSTSVOLLQRLYDPVOGOVLFDGVYAKELNOMLSQAIYQOEVLNCSIAEMI 1049  
Db 1081 SCCGKSTSVSLERYDPEGEVLVDGLSVRNLTQWMAQNGIYQOEILDCSISGDI 1140  
QY 1050 AYGDSSRVPLDEIRKANAANAHISFIEGLPEKYNTOVGLKGAOLSGGQKORLAIRALL 1109  
Db 1141 AYGDNRKRYQGEITFAKKEAMHISFIESLTDKYNTRVGDKTQLSGGQKORLAIRALI 1200  
QY 1110 OKPKTLLEDEATSDLNDESEKYVQNALDKARTGRCLVTHRLSALONADLIIVLHNGKI 1169  
Db 1201 RKPKTLLEDEATSDLTESEKYYQEALDKARMGRTCIYIAHRLSTIONADKIAVIQNGKV 1260  
QY 1170 KRGQTHQELLRNRDIFYKLYVNMQ 1192  
Db 1261 VEOGTHQQLQKLVYFSLVTIO 1283

RESULT 9  
148123  
P:glycoprotein isoform III - Chinese hamster  
C:Species: Crictetus griseus (Chinese hamster)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 02-Feb-2001  
C:Accession: 148123  
R:Endicott, J.A.; Sarang, F.; Ling, Y.  
DNA Seq. 2, 89-101, 1991  
A:Title: Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene family  
A:Reference number: 148121; MUID:92135896  
A:Accession: 148123  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1281 <RES>  
A:Cross-references: GB:M60042; NID:9191168; PIDN:AAA68885.1; PID:9191169  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop  
F:412-606/Domain: ATP-binding cassette homology <ABC1>  
F:429-436/Region: nucleotide-binding motif A (P-loop)  
F:1054-1250/Domain: ATP-binding cassette homology <ABC2>  
F:1071-1078/Region: nucleotide-binding motif A (P-loop)

Matches 669; Conservative 239; Mismatches 273; Indels 51; Gaps 9;  
QY 1 MTLGLIASLVNACPLMPLVYGEISDNLSGCLVQNTNYSF-----FRLTL----- 47  
Db 57 MLGLTMAIAHSGSLPLMIVGEMTDKFN-----NAGNSLPVNFSLSMINPGRILE 110  
QY 48 -----YVGVGAALLFGYIOISLWITTAARQTKRKRFKQFHSVLADQIGMFSDCI 99  
Db 111 EEMTKRATYYISGLGGVLAAYITQVSWFTLAAGROTKRKRONFPAHLQDEGMEDIKET 170  
QY 100 GELNTRMT-DIDKISDIDGKTALLFQNNSTFSGIAGLVGKMLTTLVTSPLIMAS 158  
Db 171 TELNRLTLDIDKISGIDGKVMFQAVATFPAGTIVGFIQKMLTIVMAISPLIGLS 230  
QY 159 AAACSRMTISLTSKELSAVSKAGAAVEVLSIRIVYAFRAOEKELQRTONLKADKFG 218  
Db 231 AAVWAKIISTFSDKEIAAAYAKAGAAVEELGAIIRVYIAGGONKELEERYOKHLENKAKIG 290  
QY 219 IKRTIASKVSIGAVYFPMGTGLAFWVGTSLINCEPEYTIQTVLAFVFSYHSYCIG 278  
Db 291 IKKAIISANISMGIAFLIYALAYALAWGSTLVISKE--YTTGNMAYTFEFLIGAFSVG 348  
QY 279 AAVPHRETPAIRGAAFHIFQVIDKKPSIDNFTAGYKPESTIGTVEFKVNSFNPSPRS 338  
Db 349 QAPCPDAPANNRGAAYVFDIIDNNPKIDSFSEGRHKPDSIKGNLDFSDVHFSPSRAN 408  
QY 339 IKTLGLNLRIRKSGETVALVGLNGSGKSTVYOLLQRLYDPDGFIVWENDIRALNRYH 398  
Db 409 IKTLGLNLRKVSQGVVALVGNSSGCKTTLQRLYPTGTGTSIDQDIRNFVRL 468  
QY 399 RDHIGVSOEPLVFGFTISNNIKYGRDVTDEEMEAAREANAYDPEPKPFNTLYGE 458  
Db 469 RELIGVSOEPLVFTTIAENIRYRGANTMEIRKAYEANAHEITMLPKQFDTLVE 528  
QY 459 KGAOMSGOKORAIARALVRNPKIILDEATSDLSESKSAVOALEKASGRRTIYVA 518  
Db 529 RGAQSLSGGQKORAIARALVRNPKIILDEATSDLSEAEVQALDKAREGRTIYVA 588  
QY 519 HRLSTRSADLIYTLKQMLAEKGAHAEIMAKRGLYSLV-----MSDI----- 563  
Db 589 HRLSTRNADVIVAGFEDGYVEGSHSELMQKEGYFKLYNMQSGSLISOFEVEELSE 648  
QY 564 KRADQEMESMTYSTERKTSPLHSYKSIKSDFIKAEESTOSKET--SLPEVSLKIILK 621  
Db 649 EKRAADGMPNGMKSHIFRNS---TKSLKSSRAHHRLDVDDELDAVNPVSEKLVK 704  
QY 622 LNKPEMPYVLTGLASVINGVHPVFSIIFAKIITMFGNNDKTKLKHAEIYSMFEVLG 681  
Db 705 LNKTEMPYFVVGCAIVAGALQPAISILISEMAIFGFGDDAVKQKCNLSVFLG 764  
QY 682 VICEVSYEMOGLFYRAGEIITMRLRHLAFKAMLYODIAMDPEKENSGGLTTLIADIA 741  
Db 765 VLSFTEFFQGFQFGKAGEIITLRLSMAFKAMLRODMSPFDYKRNKSTALSTRLTDA 824  
QY 742 OIOGATGSRIGVLTQNTNMGLSVIISFYEMEMTFLIISAPVAVTGMETGAMTGPA 801  
Db 825 QVQGATGFLALAOQNTAMLGITIIISFYGMQLTLISVVPFIIVASIVKMKMLAGNA 884  
QY 802 NKDKQELKHAGRIATEALENIRTIYSLTRKAFQOMYEMLOTOHRNTSKAQIIGSCYA 861  
Db 885 KRDKKLEAAGKRIATEALENIRTVASLTQERKESNYVEKLEHPYNSVQMAHIYGTFS 944  
QY 862 FSHAFIYFAIYAAGFRGAVLIQAGRMTPCEGMFYFAIAYGAMAIGKTLVLAPEYSKAS 921  
Db 945 ISQAFYFYSYACFRGAVLIYGNHFRFDVILVFSIAYFGAVALLCHASSFADPAKAL 1004  
QY 922 GAHHLFALLEKKNIDRSQEGKKPDTCEGNLEPREVSPYPCRPDVFIILRGLSLSTER 981  
Db 1005 SAHHLFSLEROPLLDISYSGEGLMPDKFGSVTFNENVNPTPRAMPVQGLSLSTLVKKG 1064  
QY 982 KTVAFVSSGSGCKSTSVOLLQRLYDPVOGOVLFDGVYAKELNOMLSQAIYQOEVL 1041  
Db 1065 QTLALVSSGSGCKSTVOLLERYDPEGMAGTVLLDGEAKKLNIQWRAQGLIYSGEPVLF 1124











Db 837 KAMLGODIGMEDDLRNNPGLVLTTRLATDASQVQATGSGVGMVNSFTNIIALLIAFFP 896  
Qy 772 GHEMFELLSTAPVLAATGMETAMTGKANKDQKELKAGIATDALENITIVSLTE 831  
Db 897 SMKSLITITFEPLALSGAATKMLTGFASQDKQALEMAQGITSELSNITVAGIGE 956  
Qy 832 KAFQEMEEMLOTOHRNTSKKAOIGSCYAFSHAFYFAAGFRGVALIOAGRMTPEG 891  
Db 957 GRTKAFELVELOTSKRTAVRKANITIGLCPAFSOGIAFLANSAYKRGGLIAVEGIGFSEH 1016  
Qy 892 MEYFETAAYGAMAIGKTLVLAPEXSKASGAHLFALLEKRPNDISRSQEKRPDCEG 951  
Db 1017 VERVVSVLSTATAGFRFSYPSYAKAKISARFEQLDRKRPINVSSEAGEKMDNFCQ 1076  
Qy 952 NLEFEVEFFYCRPDVFLRLSLSTIERGKTVAVGVSSGCKRSVOLLOLRYPVOG 1011  
Db 1077 KIDFIDCKFTYPRPDIOVLNLGSVNPQGLAVSGSSGCKSTSIQLEFRYDPOCT 1136  
Qy 1012 VLFDDVAKELNVQMLRSOIAIVPOEVLNCSIAENIAYGNSRVPPLDEIKKANMAN 1071  
Db 1137 VMIDHDSKKNVIOQLRSKIGIVSQEPVLFDCSINDNITKGNTEISVERAIAAKQO 1196  
Qy 1072 IHSFTEGPEKYNTOVGKGAQLSGGQKORAIARALLQKPKILLDEATSAINDSEKY 1131  
Db 1197 LHDVMSIPEKETNVGIGSQLSRGEKORAIARAIYVDRPKILLDEATSAIDTESEKT 1256  
Qy 1132 VOHALDKARTGTCVTVHRLSAIONADLIYVHLHGKIKEGTHOELRNDRIFKLV 1189  
Db 1257 VOTALDKAREGTCIVIAHRLSTIONSIDIIVAVSOGVVIETKGTHERKMAQKAYYKLV 1314

RESULT 14  
T19982  
hypothetical protein C47A10.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000  
C:Accession: T19982; T24202  
R: Basham, V.  
submitted to the EMBL Data Library, November 1996  
A: Reference number: Z19207  
A: Accession: T19982  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-1294 <M12>  
A: Cross-references: EMBL:Z81484; PIDN: CAB03973.1; GSPDB: GN00023; CESP: C47A10.1  
R: Basham, V.  
submitted to the EMBL Data Library, April 1997  
A: Reference number: Z19853  
A: Accession: T24202  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-1294 <M12>  
A: Cross-references: EMBL:Z93782; PIDN: CAB07855.1; GSPDB: GN00023; CESP: C47A10.1  
A: Experimental source: clone R12G8  
C: Geneticals:  
A: Gene: CESP: C47A10.1  
A: Map position: 5  
A: Introns: 15/2; 67/3; 123/3; 168/2; 509/1; 649/3; 826/2; 1089/3; 1158/3; 1250/1  
C: Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 42.1%; Score 2529; DB 2; Length 1294;  
Best Local Similarity 43.3%; Pred. No. 3.2e-148;  
Matches 533; Conservative 224; Mismatches 430; Indels 44; Gaps 10;

Qy 1 MILGLIASLVNACGLPMLVLVGEKSDNIS--GCLVQTN-----TYSFRL 45  
Db 46 IAVGVIGCATVGLPLMSIIGVNSQNFVTLTGTLFDLSTNASTSEKAAARAEFSHEVN 105  
Qy 46 TLYVYGIVAAALIFGYIOISLWITTAARQTKRIKOFHFSVAODIGFDSGIDGLNTR 105  
Db 106 CLKYVYLOGGIFAAAGFLQASCFMVICERLSNFRROFFHSYAROEIAWDKNTSGTSLSK 165

Qy 106 MTD-IDKISDGIKDTALLFQNMSTFSGIAGVAVGKMLTVLTSTPLIMASAAACSR 164  
Db 166 LFDNLERVEGEGTDKGLAFQMAOFTGGAAVFTYDMTLTLMKSLSPMMICGLFLK 225  
Qy 165 NVISLTSKELASVSKAGAAVEVLSIRTVIAFRAOEKELORYTOULKAKDFGIRRTIA 224  
Db 226 LATATATKKAQYAAVAGGIAEVLTSIRTVIANNGEYDEKREDALEHGKTKGKISFL 285  
Qy 225 SKVSLGAVYFPMNGTYGLAFWGTSLILNGEPGYTIGVLAVFSSYIHSYCIGAAPH 284  
Db 286 IGAGLASFEVLIYASCYCLAFWGTNTFVYSGR--LESGLVLTVEFSVMGSMALGQAGOF 343  
Qy 285 EHFATARGAAPHFOYIDKPSIDNFTAGYRPESTEGYEEFNANSPNPSRSITLNG 344  
Db 344 ATGTALGAASALYEVIDRIPETIDASTBQPSKISGRISVAKVETPTPRADVILKG 403  
Qy 345 LMLRIKSGEFTVALVGLNGSGKSTVVOLORLYDPDGFIVDENDIRALNVRRHIGV 404  
Db 404 VSLDAQPGQTVAVLGGSGCKSTIIDLQRFYRPAQGLIIDDPIEDFNIRKTLROLGV 463  
Qy 405 VSOEPLVFGTTISNNIKYGRDVTDEEMERAREANAYDIMEFPNKFNTLVGEKAQMS 464  
Db 464 VSOEPLVFTSIEQNTIRYGRSDVSDEDIAALKEANAADFIKTFPEGLMTLVGDRGQMS 523  
Qy 465 GQOKRIATARALVNRPKILLDEATSALDSEKSAVOALEKASGRITTVIAHRLSTI 524  
Db 524 GQOKRIATARALVNRPKILLDEATSALDSEKSIYOSALENASGRITTVIAHRLSTV 583  
Qy 525 RSADLIVTLKDGMLAEKGAHAEIMARKGLYSLVMSQ-----DIKADROMESMTYTE 578  
Db 584 RNADKIIVKAGQVMEVGHETLIEQKGLYHELVHQAQVADVDRKKKEAEKRRMSRQIS 643  
Qy 579 RRTNSLPLHSVSKISDFIDK-----AEESTOSKEISLPE----VSLKLIKLNK 624  
Db 644 QKRGSV---NEKTOQESQVDEKPKAPAPAEAKEREIKRLKELEBEGAVKANLEKILRYAR 700  
Qy 625 PEKPFVVLGTSLASVNGYVAPVSIIFAKITTFMGNNDKTTLKHAETISMTFVILGVC 684  
Db 701 PEWITVFALIAALIOGAVPAFSLFESQIINFNSPDRQMKDGHFAMLMVLVLAAYV 760  
Qy 685 FVSYFMQGLFYGRAGELTMRRLHFLAKMLYODIMFDEKNSGTGLTTLIDIAQIO 744  
Db 761 GTSMLFQCSLFGVAAERLIRIRISKYRNVLRDADATYFDMKPSRGRITTRLATDAPNR 820  
Qy 745 GATGSRIGVLTQATNMGLSVIISFTYGEKTEPLILISAPVLAVTGMETAAATGFAKMD 804  
Db 821 SAILYRLSGIFENAIASVGGGLIAFYGMQMAFLVMAIFPFMAVGOALMMKYHGSATSD 880  
Qy 805 KQELKAGKIATDALENIRIVSLTREKAFQOMYEMLQTOHNTSKKAOITSCAFSH 864  
Db 881 AKEMENAGKTAMAEINIRIVQALTQTKLYNIFCSLDPHOGNISKALIRDLITGFAN 940  
Qy 865 AFIFYAAGFRGAVLI--QAGRMTPEGMFIYFTAIAGCAMAIGTVLVAPEYSAKSG 922  
Db 941 SIOFTTYAAAFRGLFLIDPKNVLMEDENVLRKLPALISFSGITIGRAASFPEYIKATPA 1000  
Qy 923 AAILFALLEKRPNDISRSQEGKRPDCEGNLEFRFVSFFYCRPDVFTILGSLSTIERGK 982  
Db 1001 AGCIFMMLDEEPRIDQMTSGTYRQ--LSGEVKLANKYFFRRPERPAPVILQGLVHWKPGQ 1059  
Qy 983 TVAFVSSGGCKSTSVQLORLYDPVOGOVLPEDGVDAKELNVQMLRSOIAIVPOEVLN 1042  
Db 1060 TIALVPSGCGKSTVSLIERLTDPLEGATVYNNNDLRQMPNPKRLKRIHIAVSOEPLTFD 1119  
Qy 1043 CSTAENIAGVADNSRVPLDEIKRANAANTHSRLEGPKEYNTOVGKGAQSGGQKORL 1102  
Db 1120 TSTRINIVTGLQGEYTHQDIETACSAKNTIKHIDELPDSEYERVGEKGLQSGGQKORI 1179  
Qy 1103 AIALALQKPKILLDEATSALDNDSEKVVQALDKARTGTCVTVHRLSAIONADLIY 1162  
Db 1180 AIALALIRNPKILLDEATSALDTESEKQVVALDAAKORGTIIVAHRLSTIVNMGCM 1239



---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:32:48 ; Search time 24.77 Seconds

(without alignments) 1768.852 Million cell updates/sec

Title: US-09-873-409-6

Perfect score: 6012  
Sequence: 1 MIIIGILASLVNGACLPMLPL.....QELLRNDRIFKLVNQSVQ 1195

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3574	59.4	1276	1	MOR1_CRIGR
2	3554	59.1	1280	1	MOR1_HUMAN
3	3529	58.7	1276	1	MDR2_CRIGR
4	3526.5	58.7	1276	1	MDR1_MOUSE
5	3520.5	58.6	1276	1	MDR3_MOUSE
6	3504.5	58.3	1279	1	MDR3_HUMAN
7	3480.5	57.9	1276	1	MDR2_MOUSE
8	3476.5	57.8	1277	1	MDR1_RAT
9	3441.5	57.2	1281	1	MDR3_CRIGR
10	3428.5	57.0	1278	1	MDR2_RAT
11	2868	47.7	1321	1	AB11_HUMAN
12	2839	47.2	1321	1	AB11_RAT
13	2834	47.1	1321	1	AB11_MOUSE
14	2765	46.0	1321	1	AB11_RABIT
15	2326	38.7	1321	1	MDR1_CAEL
16	2308.5	38.4	1302	1	MDR4_DROME
17	2215	36.8	1302	1	MDR5_DROME
18	2135.5	35.5	1254	1	MDR3_CAEL
19	2107.5	35.1	1362	1	PMO1_SCHPO
20	1825	30.4	1480	1	MDR1_LEIEN
21	1479	24.6	1419	1	MDR1_PLAIF
22	1252.5	20.8	1290	1	STRE5_YEAST
23	1112	18.5	1336	1	MAM1_SCHPO
24	952	15.8	1323	1	HST6_CANAL
25	844	14.0	735	1	ABC8_HUMAN
26	773	12.9	820	1	MDL2_YEAST
27	752	12.5	598	1	Y288_THEMA
28	742	12.3	703	1	TAP2_RAT
29	739.5	12.3	695	1	MDL1_YEAST
30	738	12.3	582	1	MSBA_ECOLI
31	731.5	12.2	685	1	MDL1_CANAL
32	726	12.2	726	1	YFX9_SCHPO
33	713	11.9	702	1	TAP2_MOUSE

## ALIGNMENTS

RESULT	ID	MDR1_CRIGR	STANDARD:	PRT: 1276 AA.	
AC	P21448;	01-MAY-1991 (Rel. 18, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).				
GN	ABCB1 OR PGX1 OR PGPI.				
OS	Cricetulus griseus (Chinese hamster).				
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;				
OC	Cricetulus.				
OX	NCBI_TaxID-10029;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-92135896; PubMed-1685679;				
RA	Endicott J.A., Sarangl F., Ling V.;				
RT	*Complete cDNA sequences encoding the Chinese hamster P-glycoprotein				
RT	gene family.*;				
RL	DNA Seq. 2:89-101(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-91154265; PubMed-1671863;				
RA	Devine S.E., Hussain A., Davide J.P., Melera P.W.;				
RT	*Full length and alternatively spliced pppl transcripts in multidrug-				
RT	resistant Chinese hamster lung cells.*;				
RL	J. Biol. Chem. 266:4545-4555(1991).				
RN	[3]				
RP	SEQUENCE OF 706-1276 FROM N.A.				
RX	MEDLINE-88122132; PubMed-2893255;				
RA	Endicott J.A., Juranaka P.F., Sarangl F., Gerlach J.H., Deuchars K.L.,				
RT	Ling V.;				
RT	*Simultaneous expression of two P-glycoprotein genes in				
RT	drug-sensitive Chinese hamster ovary cells.*;				
RL	Mol. Cell. Biol. 7:4075-4081(1987).				
CC	-1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED				
CC	DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC	-1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT				
CC	CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY				
CC	CANNOT.				
CC	-1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY				
CC	(ABC TRANSPORTERS). MDR SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
CC	EMBL; M60040; AAA68883.1; -				
CC	EMBL; M59253; AAA37004.1; -				
CC	EMBL; M17897; AAA37006.1; -				

DR PIR; A38696; DVHVIC.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001140; ABC\_transport\_cmem.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR POSTITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
 KW Multigene family.  
 FT DOMAIN 1 50 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 51 71 POTENTIAL.  
 FT TRANSMEM 117 137 POTENTIAL.  
 FT TRANSMEM 186 206 POTENTIAL.  
 FT TRANSMEM 213 233 POTENTIAL.  
 FT TRANSMEM 294 314 POTENTIAL.  
 FT TRANSMEM 323 343 POTENTIAL.  
 FT DOMAIN 344 707 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 708 728 POTENTIAL.  
 FT TRANSMEM 754 774 POTENTIAL.  
 FT TRANSMEM 830 850 POTENTIAL.  
 FT TRANSMEM 851 871 POTENTIAL.  
 FT TRANSMEM 934 954 POTENTIAL.  
 FT TRANSMEM 971 991 POTENTIAL.  
 FT DOMAIN 992 1276 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 424 431 ATP (POTENTIAL).  
 FT NP\_BIND 1067 1074 ATP (POTENTIAL).  
 FT REPEAT 1 635  
 FT REPEAT 636 1276  
 FT CONFLICT 338 339  
 FT SEQUENCE 1276 AA; 140925 MW; 44F3F92A1864DEF CRC64;

Query Match 59.4%; Score 3574; DB 1; Length 1276;  
 Best Local Similarity 55.2%; Pred. No. 1.4e-200;  
 Matches 679; Conservative 248; Mismatches 259; Indels 44; Gaps 8;

QY 1 MILGLSLVNGACPLMPLVLGEMSDNLISGLIVQTN----- 38  
 DB 50 MLVGLLAIVHGVALPLMLVFGDMTDSFASGNIPTNATNNOVNASDIFGKLEEMT 109  
 QY 39 TYSEFRLTYVIGVALIFGYIDISLITTAARQTRKIRKOFHSYLAODIGWFSCD 98  
 DB 110 TYAV-----YTGAGVGLIYAIQVSWFCLAGRQIKHQKFEHALMNGEIGFVDHD 164  
 QY 99 IGEINTRT-DIDKISDGIKIALIFONMSTFSLAGLVGKGLLVLTSLPLMA 157  
 DB 165 VGEINTRLTDVSKINEIGDKIGFQAMATFFGGLIIGTRGKLLVLIATSPVLGL 224  
 QY 158 SAACSRVVISLTSKELSAYSKAGAAVEEVLSTRTVIAFRAOEELGRYQNLKADKF 217  
 DB 225 SAGIWAATLSTFKELDAIKAGAAVEEVLATRTVIAFGQKELEKYNNELEAKRL 284  
 QY 218 GIKRTIASKVSLGAVYFPMNGTYGLAFWYGTSLINGEPYITGLVAFVSHSYCI 277  
 DB 285 GIKKATITANISMGAAFLIYASALAFWYGTSLVSK--YISGVLYVFAVLIGARSI 342  
 QY 278 GAAPVHFTFALRGAARHITQVLDKPSIDNFTAGYKPSIECTVEKNVSENPSPRP 337  
 DB 343 GQASPNIETAFNARGAAEIFNIIDNKPSIDFSKNGYKPDINKNLEFKNIHFSPSRK 402  
 QY 338 SIKITKGLNTRKSGETALVGLNGSGKSTVOLLQRLYDPDGIIMVDENDIRALNRRH 397  
 DB 403 DVQIILKGLNLVQSGOTVALVNGSGKSTVOLLQRLYDPDGIIMVDENDIRALNRRH 462  
 QY 398 YRDIGVSOEPLVGGTITSNIKIKYGRDVTDEEMERARAEANADFTIMEFNKNFTLVG 457  
 DB 463 LREITGVSOEPLVFAITIAENIRKRENVYMDIEKAVKENADFTIMKPKHCDITLVG 522  
 QY 458 EKGAMSGGOKORIALARALVNRPKILLDEATSLDSEKSAVOALEKASKGRTIV 517  
 DB 523 ERGAQLSGGOKORIALARALVNRPKILLDEATSLDTESEAVVQALDKAREGRTIIV 582

QY 518 AHRISTIRADLIYTLKDGMLAEKGAHAEMLAKRGLYSLVMSQ-----DIKRD 567  
 DB 583 AHRISTVNAIIAGFDGCVIVEQGNHEELKREKGYIKYKIVMTQVAGNEIELGNVSGSK 642  
 QY 568 EOMESMTYSTERKTNLS--PLHSYKSTKSDFIIDAEEST-OSKEISLPEVSLIKLIK 624  
 DB 643 NEIDLMDSSKDSASLIRRRSTRSIRGPHDQDKLSTKALDEDPVIFSRILKINS 702  
 QY 625 PEMPVVGLTASVYNGVHVHVFSTIFAKITMF-GNNDKTKLKDAIYIMIPITIGVI 683  
 DB 703 SEMPVFVVGICAIYNGALQPAFSTIFSKVVGVEFRNTDDETKRDSLSFLPLIGVI 762  
 QY 684 CEFVSFMGLYFGRAGELITRLRLRLAEKAMLYODIAMEFDEKENSTGGLITLAIIDAOI 743  
 DB 763 SFITFFLOGFTFGKAGELITRLRLRYMFKSMRLQDVSPFMDPKNTGTLTIRLANDAOV 822  
 QY 744 QGATSGRIGVLQNTNNGLSVIISFIYEWENTFLLSIAPLAVTGMETPAAMTGFAK 803  
 DB 823 KGATGARLAVITQNTANLGTGIIISLIYQWDTLLLAIIVPILAIAGVEMKMLSGQALK 882  
 QY 804 DKQELKHAKTATELENIIRTVSLTRKAFQWYEMLOQOHRNTSKAODIGSCVAFS 863  
 DB 883 DKKELEGSKTATELENIIRTVSLTRQKFNMTAOSLQIPYRNALKKAHVFITFSFT 942  
 QY 864 HAFIYFAAGRFEGAYLQAGRMTPCEGMFVFTAIAGAMAGICTLVLAPEYSKASGA 923  
 DB 943 QAMATFSTYACGRFAYLYVARLMTFENVLVFSIYVGAAMAVGVSSFADYNAKAYSA 1002  
 QY 924 AHLFALLEKKNPIDSRSQEGKPDTCENLEFEVSFFPCRPDVFILRGLSLSTERKT 983  
 DB 1003 SHIIMITEKVPISIDSYSGKLKPNLTLEGNVAFENVEVFNPTDPIDPLVGLNLEVKKQOT 1062  
 QY 984 VAFVSSGCGKSTSVOLLQRLYDPVQGVLPDGVAKELNQMLSOAIYDQEVLEFNC 1043  
 DB 1063 LALVSSCGKSTSVOLLERFYPDPAAGTVFIDGKVNQMLQWMLAHGLIVSOEPLIFDC 1122  
 QY 1044 STAEINAYGDSNRVPLDEIKENANANISFIEGLPEKYNTQVGLKAGLQSGGOKORIA 1103  
 DB 1123 STAEINAYGDSNRVSOEIERAKENANIHOPIESLPKYNRVYDKGTQLSGGOKORIA 1182  
 QY 1104 IARALLQPKILLDEATSLDNDSEKVVQHALDKARTGRTCLVTVHRLSALQNDLIV 1163  
 DB 1183 IARALVRPHILLDEATSLDTESEKVVQEALDKARREGRTIVIAHRLSTQNDLIV 1242  
 QY 1164 LHNGKIKQCGTHQELLRNRDIYFKLVNOS 1193  
 DB 1243 IONGKVKHGTQQLLAQKGYFSMVQVA 1272  
 RESULT 2  
 MDRL\_HUMAN STANDARD; PRT; 1280 AA.  
 ID MDRL\_HUMAN Q12755; Q14812;  
 AC P08183; Q12755; Q14812;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MUTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).  
 GN ABCB1 OR PGY1 OR MDRL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87028230; PubMed=2876781;  
 RA Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,  
 RA Robinson J.B.;  
 RT "Internal duplication and homology with bacterial transport proteins  
 RT in the mdrl (P-glycoprotein) gene from multidrug-resistant human  
 RT cells.";  
 RL Cell 47:381-389(1986).



Db 349 SPSTEAANMAGAYEIKFIKIDNKPSTDSYSKSHKPDNIGNEFRNVHSPYSRKEVK 408  
401 ILKGNTRIKSGEVALVGLNGSGKSPVOLLQRLYDDPGEIMVDENDIALVNRHD 400  
409 ILKLNINAVSGQVVALVGNNGCKSTTVOLMORLYDPTCEMVSVDGODITTVNFRRE 468  
401 HIGVSOEPLVFGTTISNNIKYGRDVTDEMERAREANADEFIMEPNKENTLVGSKG 460  
469 IIGVSGQEPVLFATTINENIRINGRENTVMDIEKAVKEANVDFIMKPKHFDLVGSRG 528  
461 AQMSGGOKORAIARALVNRKIIILDEATSDLESSEKSAVQALAKASKRTTIVAAH 520  
529 AQMSGGOKORAIARALVNRKIIILDEATSDLESSEKSAVQALAKASKRTTIVAAH 588  
521 LSTRSADLYTLKDMGLAEKGAHAEIMAKRGVLSVMSQ-----DIKKADEMESMT 574  
589 LSTRSADLYTLKDMGLAEKGAHAEIMAKRGVLSVMSQ-----DIKKADEMESMT 648  
575 YSTERKNTSLPLHSV-----KSIR-SDPDKAEESTOSKEISLPEVSLKILKLNKEW 627  
649 DALEMSNSDSKSLIRKSTRSRVRSQAORSLSTKALDESIPVPSFWRMKLNLEW 708  
628 PFVVLGLASVLTAVHVSIIIRAKIITMGN-NDKTTLKHAEIYSMTIVILGICFV 686  
709 PYFVAGFCAIINGLOPAFAIIFSKIGVTRIDDPETKONSMLFSLFALGIIISFI 768  
687 SYFMGLFYGAGELIMRLRHLAFKAMLYODIAMPEKENGSTGLTIIILIDAOIGA 746  
769 TFFIAGFTFGAGELLTKRLRYMFRSLRQDVSWFDDPKTGLTLRLNDAQVGA 828  
747 TGSRIQVLTQATNMGSLVSIISFYGEMTFLILSIAPVLAVTGMETAAITGAFANKDO 806  
829 IGSRLAVTQIANMGIIISFYGMQLTLLAIYPIIAGVVEKMLSGQALDKDK 888  
807 ELKAGIATATLAINIRIYSLTREKAEOMYEMLOQHNKTSKKAQIISCAFAHAF 866  
889 ELEAGKATATEIENFRIVSLTQOKFEHMYAOSLOVYRNSLRKAIIFITPSFTOM 948  
867 IYFAYAGFRGAYLIOGRMTPEGMFVFAIYVAGAMAKTIVLAPEYSKAKSAHL 926  
949 MYFSYACFRGAYLVIAHKLMSFEDVILVFSAYVFGMAVQVSSAFDYAKAKISAHI 1008  
927 FALLEKPNIDRSQEGKPDCEGNEFRVSFEYPCRPDVFILRGISLIERGKYAF 986  
1009 IMIEKTPILDSYSTEGIMPTLEGVTFGEVNYPTRPDIYPVLOGLSLEKKGQTLAL 1068  
987 VGSSGCKSTVOLQRLYDVOGVLFDGVDAKELNQVLRSAIAYPOEPVLFNCIA 1046  
1069 VGSSGCKSTVOLQRLYDVOGVLFDGVDAKELNQVLRSAIAYPOEPVLFNCIA 1128  
1047 ENIYAGNSRVVPLDEIKEAANANINHSIFGLEPKYNTQVGLKGAOLSGOKORAIAR 1106  
1129 ENIYAGNSRVVPLDEIKEAANINHSIFGLEPKYNTQVGLKGAOLSGOKORAIAR 1188-  
1107 ALLOKPKILLDEATSDLESSEKSAVQALAKASKRTTIVAAH 1166  
1189 ALVQPHILLDEATSDLESSEKSAVQALAKASKRTTIVAAH 1248  
1167 GKKEGTHOELLNROIFKLVNAQS 1193  
1249 GRAKEGHQOQLAQGIYFSMVSQA 1275

RESULT 3  
MDR2\_CRIGR  
ID MDR2\_CRIGR STANDARD; PRT: 1276 AA.  
AC P21449;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).  
GN PGY2 OR PGP2.

OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92135896; PubMed=1685679;  
RA Endicott J.A., Sarangi F., Ling V.;  
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein  
RT gene family";  
RL DNA Seq. 2:89-101(1991).  
RN [2]  
RP SEQUENCE OF 622-1276 FROM N.A.  
RX MEDLINE=88122132; PubMed=289325;  
RA Endicott J.A., Juranaka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,  
RA Ling V.;  
RT "Simultaneous expression of two P-glycoprotein genes in  
RT drug-sensitive Chinese hamster ovary cells.";  
RL Mol. Cell. Biol. 7:4075-4081(1987).  
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC CAPABILITIES: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT  
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY  
CC CANNOT.  
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; M60041; AAA68884.1; -;  
DR EMBL; M17896; AAA37007.1; -;  
DR PIR; B27126; DVHY2C.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003439; ABC\_transport.  
DR InterPro: IPR001140; ABC\_transport\_tmem.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam; PF00664; ABC\_membrane; 2.  
DR Pfam; PF00005; ABC\_tran; 2.  
DR SMART; SM00382; AAA; 2.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
KW Multigene family.  
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).  
FT TRANSSEM 52 72 POTENTIAL.  
FT TRANSSEM 119 139 POTENTIAL.  
FT TRANSSEM 188 208 POTENTIAL.  
FT TRANSSEM 215 235 POTENTIAL.  
FT TRANSSEM 296 316 POTENTIAL.  
FT TRANSSEM 325 345 POTENTIAL.  
FT TRANSSEM 346 708 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 709 729 POTENTIAL.  
FT TRANSSEM 755 775 POTENTIAL.  
FT TRANSSEM 831 851 POTENTIAL.  
FT TRANSSEM 852 872 POTENTIAL.  
FT TRANSSEM 935 955 POTENTIAL.  
FT TRANSSEM 972 992 POTENTIAL.  
FT DOMAIN 993 1276 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 426 433 ATP (POTENTIAL).  
FT NP\_BIND 1068 1075 ATP (POTENTIAL).  
FT REPEAT 1 635  
FT REPEAT 636 1276  
SQ SEQUENCE 1276 AA; 141057 MW; 5096B1385628812D CRC64;

Query Match

58.7%; Score 3529; DB 1; Length 1276;



Best Local Similarity 55.7%, Pred. No. 5,8e-198;  
Matches 685; Conservative 232; Mismatches 266; Indels 46; Gaps 9;

```
OY 1 MILDGLAVNGACLPMPVLYGEMSDN-----LISGLCVO-T 37
Db 51 MVLGLAVLVHGTSLPLMLVFGNMNTDSTFKAETSIMPNMNTQSEINNTNTEISGLEEDM 110
OY 38 MYSPFRLTLVYGVIGVALIFGVIQISLWITTAROTTRIKRQPFHSLADIGMFOSC 97
Db 111 ATYAI-----YTTGIGAGVLYVAYIYQVSPWCLAAGRQINKIKQFFHAIWNOEIGFDVH 165
OY 98 DIGELNMT-DIDKISDGIKDKALLFQNMSTFISGLAVGLWKMCKLTLVLTSPIM 156
Db 166 DIGELNMTLTDVSKINDGIDGKIMFQSIATFLAFLIVGISGKMLVLYLVASPLIG 225
OY 157 ASAAACSMYVLSLSKELSAVSKAGAAVEVLSIRIVYIAFRAQEKLORYQNTKDARD 216
Db 226 LSSAMMAKVLTSFTNKELOAAKAGAAVEVLAIRTVIAFGQNKLELERYKNMLEEAKN 285
OY 217 FGIKRTIAVKVSLGAVYFPMNGTGLAFMYGSLTLNGPQGTIGTVLAVFESVHSYC 276
Db 286 VGKKAIVANISIGIAYLVASVYALAWYGSVLASNE--YSGVQVLFVFSILFETFS 343
OY 277 IGAAPVHEFETAIARGAAFHIFQVYIDKRRPSIDNESTAGYKPPESIGTVEFKNVSEYPSR 336
Db 344 IGHIAVNEVEFANMARGAAVEIFKIIDNEPSIDSFSTOGKHPDSVGNLEFKNVHVSYSR 403
OY 337 PSIKTLKGLNRIKSGEVALVGLNGSKSVYOLLORYLDPDGFIMVENDIRALNV 396
Db 404 SGIKTLKGLNRIKSGEVALVGLNGSKSVYOLLORYLDPDGFIMVENDIRALNV 463
OY 397 HYRHOIGVAVSOPEVLEGGTISNNIKYGRDVTDEMERARANVYDFEMEPNKNTLV 456
Db 464 YLREITIGVSOPEVLEGGTISNNIKYGRDVTDEMERARANVYDFEMEPNKNTLV 523
OY 457 GERGAOMSGGQKORIAIARALVRNPKILIDEATSALDSESKSAVQAALAKASKGRTTV 516
Db 524 GERGAOMSGGQKORIAIARALVRNPKILIDEATSALDSESKSAVQAALAKASKGRTTV 583
OY 517 VAHRLSTRASDLVLYTKDGMALAEKGAHMAELAKRGLYSLVMSQ-----DIK 564
Db 584 IAHRLSTRASDLVLYTKDGMALAEKGAHMAELAKRGLYSLVMSQ-----DIK 643
OY 565 KADEQMESMTYSTERKTSNLSPLHSVKS-KSPDIDKAEESTQSKESLPEVSLTKLTKLN 623
Db 644 QSDTIASELTF-SEEFKSPSVKSTGRSICGSODQERRSVKKAQEDVLYVFWGILTKLN 702
OY 624 KPEMPFVVLGTLASVLTNGTVHPFSLIFAKITMF-GNNDKTKLKHDAEYSMIFVILGV 682
Db 703 ITEMVLYVGVLCANVINGCMQVPSIVSGIIGVTRDDPRTKOONCMFLSFLVYMGH 762
OY 683 ICEVSYEMOGLFVYGRAGELITMLRHLHAKAMLYODIANFDEKENSTGILTTILADIAO 742
Db 763 ICEVSYEMOGLFVYGRAGELITMLRHLHAKAMLYODIANFDEKENSTGILTTILADIAO 822
OY 743 IOGAGSRLGVLTONATNMGSLVTSIFYGEMTFLILISAIVLAVTGMIEFRAAMGFAN 802
Db 823 VKGAASSRRLAGTTONVANNGIILISLYVGWQLTLLVLAIPILILISGMKEKVVLSGQAL 882
OY 803 KDKQELKAGKATATALENIRTVLSLREKAFQOMYEEMLOTOHRTSKAOIIGSCYAF 862
Db 883 KDKKLELVSGKATATALENIRTVLSLREKAFQOMYEEMLOTOHRTSKAOIIGSCYAF 942
OY 863 SHAFTYFAYAGFRGAYVLIQAGRMTPPEGMIFVTAIAYAGAMAGIKTVLADYESSKAGS 922
Db 943 TOAMMYFSAACFRGAYVLIQAGRMTPPEGMIFVTAIAYAGAMAGIKTVLADYESSKAGS 1002
OY 923 AAHLFALLEKRNDSRQSEKPKDTCBGNLFEVSEFFYPPORPDVFIIRGLSLSTERK 982
Db 1003 ASHILRIKMKIPSIDYSTGRGLKPMWLEGNVAFNEVEVFPYPRPDIPVLOGSLSEVKKO 1062
OY 983 TVAFVSSGCGKSTVOLLORLYDPOGOVLPDGVADAKELANQWMLRSQALAYDQEPVLN 1042
Db 1062 TVAFVSSGCGKSTVOLLORLYDPOGOVLPDGVADAKELANQWMLRSQALAYDQEPVLN 1102
```

```
Db 1063 TIALVSSGCGKSTVOLLOREFYDPNAGTVFLDKEIKOLNQMWRANHIGVQEPILPD 1122
OY 1043 CSAENIAYGDNRRVYPLEIKRANANHSIFELPEKYNVQVGLKGAOLSGGOKORL 1102
Db 1123 CSAENIAYGDNRRVYQDIERAKKANHOFIESLPDKYNTRVDDKGTQSLSGGOKORL 1182
OY 1103 ATARALVPRPHILLDEATSALDNDSEKVVQALDARFGRTRCLVYTHRLSAIONMDLY 1162
Db 1183 ATARALVPRPHILLDEATSALDNDSEKVVQALDARFGRTRCLVYTHRLSAIONMDLY 1242
OY 1163 VLNCKRIKEDGTHQELLRNDRDYFKLVNA 1191
Db 1243 VLNCKRIKEDGTHQELLRNDRDYFKLVNA 1271

RESULT 4
MDRL_MOUSE STRAND; PRT; 1276 AA.
AC P06795;
DT 01-JAN-1988 (rel. 06, last sequence update)
DE 20-AUG-2001 (rel. 40, last annotation update)
GN ABCB1 OR PGY1 OR PGY1-1 OR MDRL OR MDRLB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI-Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87028229; PubMed=3768958;
RA Gros P., Croop J., Housman D.;
RT "Mammalian multidrug resistance gene: complete cDNA sequence
RT indicates strong homology to bacterial transport proteins.";
RL Cell 47:371-380(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89367274; PubMed=2570420;
RA Raymond M., Gros P.;
RT "Mammalian multidrug-resistance gene: correlation of exon
RT organization with structural domains and duplication of an ancestral
RT gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).
RN [3]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=91042535; PubMed=2248681;
RA Raymond M., Gros P.;
RT "Cell-specific activity of cis-acting regulatory elements in the
RT promoter of the mouse multidrug resistance gene mdrl.";
RL Mol. Cell. Biol. 10:6036-6040(1990).
CC -I- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
CC LINKER DOMAIN.
CC -I- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14757; AAA79005.1; -
DR EMBL: M60348; AAA39513.1; -
DR PIR: A33719; DVM51.
DR MGI: 97568; Abcb1.
DR InterPro: IPR003593; AAA.
```

DR	InterPro: IPR003439; ABC transporter.
DR	InterPro: IPR001140; ABC transporter_tmem.
DR	InterPro: IPR001687; ATP_GTP_A.
DR	Pfam: PF00664; ABC_membrane; 2.
DR	Pfam: PF00005; ABC_tran; 2.
DR	SMART: SM00382; AAC; 2.
DR	PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR	ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW	Multigene family; Phosphorylation.
FT	DOMAIN
FT	1
FT	TRANSMEM
FT	48
FT	71
FT	TRANSMEM
FT	119
FT	139
FT	TRANSMEM
FT	199
FT	219
FT	TRANSMEM
FT	296
FT	316
FT	TRANSMEM
FT	327
FT	347
FT	DOMAIN
FT	348
FT	709
FT	TRANSMEM
FT	710
FT	730
FT	TRANSMEM
FT	755
FT	775
FT	TRANSMEM
FT	831
FT	851
FT	TRANSMEM
FT	855
FT	882
FT	TRANSMEM
FT	935
FT	955
FT	TRANSMEM
FT	966
FT	986
FT	DOMAIN
FT	987
FT	1276
FT	NP_BIND
FT	426
FT	433
FT	NP_BIND
FT	1068
FT	1075
FT	REPEAT
FT	1
FT	635
FT	REPEAT
FT	636
FT	1276
FT	CARBOHYD
FT	73
FT	73
FT	CARBOHYD
FT	91
FT	91
FT	CARBOHYD
FT	96
FT	96
FT	CARBOHYD
FT	103
FT	103
FT	SEQUENCE
FT	1276 AA; 140993 MW; 1804DFOF11BFF4E CRC64;

Query Match	58.7%;	Score 3526.5;	DB 1;	Length 1276;
Best Local Similarity	55.1%;	Pred. No. 8.1e-198;		
Matches 678;	Conservative 242;	Mismatches 262;	Indels 49;	Gaps 10

```

Db 50 MILGTLAAIHGHTLLPLMLVFGNMTDSEFTAEMSILPSTINSGSPNSTLIISMSLEE 106
Oy 40 ---YSFRLTLTYVGVGVAALLFGYIOISLMTIITAARQTKRIRKOFHSHVLAODIGWFS 96
Db 110 MAIYAY-----YTCIGAGVILVAIYQVSLWCMAAGRQIHKRQKFFHAINOEIGMDV 166
Oy 97 CDIGSLNRMF-DIDKISDGTGDKRALLFQNMSTPISGLANGVKGKMLTYLTSTPLI 155
Db 165 HDVGLNLRLLTDNDVKINDGIGDKIGMFEFOSITFFFLAFGIIIGFSGKMLTVIILAVSLP 222
Oy 156 MASAACSRMYISLTSKELSAVSKAGAAEVLSSIRFVIAFRAOEKELORYTONLQDAK 211
Db 225 GLSSLLAMKAVLTSFNKELQAYAKAGAAVEVLAIRVIAFGGOQKLEERYKNKLEBAK 288
Oy 216 DFGIKRTASKVSLGAVYFFMNGTSLGIAFWYGTSLIINGEBOYTGIVLAVFVSIHSSY 275
Db 285 NVGIKKATIASISIGIAYLLVYASALAFWYGTSLVLSNE--YSIGEVLYVFSSILGTF 342
Oy 276 CIGAAVPHFEFPAIARGAFAHFIOYIDKKPSIDNFSTRAGYRPESECTGEVKKNSFNTPS 333
Db 343 SIGHLAPPIEAFANRGAFAFELFKLIDNEPISDSTGYKPDSDIMGNLEKHNHFNPS 402
Oy 336 RPSIKILKGLNIRIKSGFEVLAIVGNSGSKSTVYOLLRQLDPDGFMDENQIDRLNV 395
Db 403 RSEVOILKGLNKKAKSGQTVLAVGNSGGCKSTTVOLMORLIDLPLEGVVSDIGODIRITNV 466
Oy 396 RHYRDHIGVNSQEPVLEFGTTISNNIKYGRDVTDEEMERARAEANAYDFIEMEPNKENTL 455
Db 463 RYLRREIIGVNSQEPVLFATTAIENIRYGRDVTDEIEMEKAVKANAYDFINKLRHQBDTL 522
Oy 456 VGEKGAOMSGGOKAIARLARLVRRPKLLIIDEATSAIDSKSNAVQALEKASKGRITI 511
Db 523 VGERAQAISGGOKAIARLARLVRRPKLLIIDEATSAIDSEKAVVQAALDKAEGRITI 582

```

QY	516	VVAHHLSLRISADLIVTLKDGMALEKGNHALLMAKRCGLIYSLVMSQDIKKADBEQMSMTY	575
Db	563	VIAHRLSVIRNADYIAGFDGVEIVIQGNHDEMEKGIYELVMTQI-TRGNLEPGNNAY	6411
QY	576	STERTNSLPLHVSYSIKSDPEIDKA-----EESTQSGEISL-----PEVSLIKILK	6212
Db	642	GSQSDTSDASELTSSEBS-KSLPIRISITSYVHRKODQERLSMKVAEDVEDPVLVSFWMLN	7000
QY	622	LNKPPEFVVLGTLASVINGTVHPVEFSTIEPAKIIITMFG-ANNKDTTLKHDEIYSMTIVIL	6800
Db	701	LNLSMPPLLVGLVLCVAVNGICQIOPFAIVAFVEIRIVGESRDDHETKRONCNLFSEFLVM	7600
QY	681	GVICVSVFMQGLFGRAGEIITMRRLRLKAMLYODIAPFDEKENSSTGLTITLADI	7400
Db	761	GLISVITVTFEQQFTFGKAGELITLRKRYMTKSMLRDIIISFDHKSSTGLTITRLASDA	8200
QY	741	AQIQGATSSRIGVLTONATNNGLSVITISFIYGMWETELIISIAFLVAVGMTIETAAITGF	8000
Db	821	SSVKCAMGARLAVVYIQNVANLIGTGIVLISLVYGMQDITLLIVITPLIYLGIIEMKLLISQ	8800
QY	801	ANKDKQELKHAGKATATELENIPTIVSLTRKFAEOMTEEMLOTOHNKTSKKAQIISCY	8600
Db	881	ALKDKQLEISGKATATEIAIENFRTIVSLTRQKFEETMAOSLOQVPRNANKKAHVSGITF	9400
QY	861	AFSHAFIPEVAAAGRFCAVYLIQAGRMPEEMFLIPFATAIAYAGKTLVLAPEYSKAK	9200
Db	941	SFTQAMMFSTPACGRFPAVYLAQOOLMEFENVMIVFSVAVFGAANAQNTSPADYAKAK	10000
QY	921	SGAHLFALLEKKPINDISRSQEGKKPDTCEGNLEFREVSEFPYPCRPDVLIRGLISIER	9800
Db	1001	VSASHIIRITEKPTRIDSYSTEGKLPTLLEGNVAFNGVQFYPTPRNIPVLQGLISLEVK	10600
QY	981	GKTVAFVSSSGCGKSTSVQLLOQLYDPVOQGYLEDGDVDAKLANQMLRSQIATVPOGFVL	10400
Db	1061	GQTLALVSSCGSGKSTVQVLLERFEDPMAGSVFLDGKEIKQLANQMLRAHLGIYSQEPIL	11200
QY	1041	FNCISIAENIAYGDNRRVYPLDEIKENAAANMHSFIEGLPERYMTQYGLKGAOLSGGQK	11000
Db	1121	FDQCSIAENIAYGDNRSRAVSHHEIYRAKAEANIHOPIFISLPKRYTRRGDKGTQLSGGQK	11800
QY	1101	RLAIARALLQPKILLDEATSAIDNSEKVVQAHLDKARTGTCVLVYHRLSAIQNADL	11600
Db	1181	RLAIARALVRPHILLDEATSAIDTESEKVVQEBALDKARGRTCIYIAHRLSTIQNADL	12400
QY	1161	IYVLIHNGKIKEGGTHQELLRRNDIYFKLYNA	1191
Db	1241	IYVLIENGKVKHEGTHQOLLAKGIFYFSWQA	1271
RESULT 5			
MDR3_MOUSE	MDR3_MOUSE	STANDARD:	PRT: 1276 AA.
AC	P21447		
DT	01-MAY-1991 (Rel. 18, Created)		
DT	01-APR-1993 (Rel. 25, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	MURFIDNG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3) (MDRLA).		
GN	ABC84 OR PGY3 OR PGF-3 OR MDR3 OR MOR1A.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10990;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=90205845; PubMed=1969610;		
RA	DeVault A., Gros P.;		
RT	"Two members of the mouse mdr gene family confer multidrug resistance		
RL	mol. Cell. Biol. 10:1652-1663(1990)."		
RN	[2]		
RX	SEQUENCE FROM N.A.		
RX	MEDLINE=90287150; PubMed=1972547;		

Seq	Sequence	1276 AA	140754 MM	75C71F33E1F58481 CRC64
50	SEQUENCE	1276 AA	140754 MM	75C71F33E1F58481 CRC64
Query Match	58.6%	Score 3520.5	DB 1	Length 1276
Best Local Similarity	54.3%	Pred. No. 1.8e-197		
Matches	668	Conservative	251	Mismatches 265; Indels 47; Gaps 8
OY	1	MILGILASLVNACPLPLMPVLIGENSDNLISGLVQTN-----T	39	
DB	50	MLVGTAAIIHGVALLPLMLIFGDMTDSFASGVNKSSTNNSEADKRAMFALEEMTT	109	
OY	40	YSFFPLTLVYGVGVAALIFGVIQISLMTITTAARQTRKIRKQFHFYSVLAODIGWDSODI	99	
DB	110	YAY-----YVTGIGAGVILVAYIOVSFPCVLAAGRDIIHRIKQFPAHIMNQGEGWDVHDV	164	
OY	100	GELNTRMT-DIDKISDGIQDITALLFQNNMSTPISGLAAGVLGKMLTIVLTSTSLIMAS	158	
DB	165	GELNTRLTDVDSKINIEGIGDKIGTFQOMATPPFGFIITGRGKMLTIVLILASVGLIS	224	
OY	159	AAASRWVSIISTKELSAVSKAGAAVEEVLISRTVIAFRAOKELORYTONLKDADQF	218	
DB	225	AGIWAIKIISFDDKELHAYAKAGAAVEELAIPTVIAFGCGKCELERYNNNLEEKRLG	284	
OY	219	IKRTIASVNSIGAVYEFMNGTGTGLAFWVGTSLINGEGYTGTVLAFVFSYHSYCG	278	
DB	285	IKKATTAISKMAAPLLIYASTALAFWGTSLVTSKE--YSIGQVLTVPFVSLIAFVS	342	
OY	279	AAVPEFETALIRGAAFIHFQVIDAKPSIDNSTAGYKPESEIEGVEEKNSFNPSPRS	338	
DB	343	QASPIIEAFANNRGAAYEVFKITIDKPSIDSTSKSGKHPNDUGLERNHIFSTPSKE	402	
OY	339	IKILGLMLRIKSGEFTVALVGLNCGSKSTVQQLRLYDPDDGFTMVENDIRALNVRY	398	
DB	403	VQILEGLMLKVSQGTVALVGNCGSKSTVQQLRLYDPDLGAVNSIDQDRIITNVRL	462	
OY	399	RDHIGVNSOEPLVEGTTISNNIKTYRDOVTDDEMRARERANAYTFIMEFPKENTLGE	458	
DB	463	RELIGVNSOEPLVFTTTIAENIRYREDVYDMEIKKAYEANAADYFIMKLPHQFDLGE	522	
OY	459	KGAOMSGGOKOIALAARLVRNPKILLIDEATSLDSEKSAVQOALAEKASGRITTIYA	518	
DB	523	KGAQSSGGOKOIALAARLVRNPKILLIDEATSLDSEAVQOALDAKARGRTTIYA	582	
OY	519	HRLSTIRSAADLVTLTKDMLAEKGAHETLMARKGLYSLVMSO-----DIKKADE	568	
DB	583	HRLSTVRNADVYAGDGVGIVGQNHDELMREKGIYFKLYMQTANGTELEGNKAEKSD	642	
OY	569	OMESKTYTTERKTNLPLHYSKSIKSDPT----DKAEBSTOSKEISLPEVSLKILKN	623	
DB	643	EIDNIDMSKSSGSSLIR--RSTFRKSICGPHDQRIKSTKEALDEDDVPAPASFWRILKN	700	
OY	624	KPEPFVVLGLTASLVNGLVHPHPSFIIFAKITIDMGN-NDKTYLKHDAEIVSMIVIIIGV	682	
DB	701	STEMPEYVGLFCALITINGLOLPASVITSKVGVVTNGPPTQONSLISLFLILGI	760	
OY	683	ICFVSYFMQGLEFYGRAGELLTMRLHHLAFKAMLYODIAMFDEKENSTGSLTTLIDIAQ	742	
DB	761	ISFTIFLQGFEGFAGELLTRKLYVMFKSLRDVSMFDDPKMTTGALTTRLRLLANDAAQ	820	
OY	743	IOGATGSRIGVLTQATNMGSLVITISFYIGWEMTFLISIAVLAVTGMIETAAATGPAN	802	
DB	821	VKGAAGSRILAVIFQNIAMIGTGIIISLIYGMQTLTLLAIVIIMAGVEKKMLSGOAL	880	
OY	803	KDKQELKAGKATATALENIRTVISLTTEKAPAEQWYEEMLOTHNNTSKKQOIGSCAF	862	
DB	881	KDKKLEESGKATATAINFRIVSLSTEQKEETMYAOSLOPYNNAMKKAHVIFITEFF	940	
OY	863	SHAFYFPAAGFREGVALIQGRMTPEGMIFVFAIYAGAMAIKTLVLAPESKASG	922	
DB	941	TOAMMYFETACFRGATLVQOGLTFENVLIVFSAIYFGMAAVQVSSFADYAKAVYS	1000	
OY	923	AAHLFALLEKRPNDISRSQEGKKPDTCEGNLEFEVSEFPCRPDVITLRLGISIERGK	982	

```

Cc Db 1001 ASHIRIRIEKTPEDISYSTOGUKPMLNEGNOVSGFVENYPTRPISIPVLOGISLEFKKGQ 1060
Cc Qy 993 TVAFVGSSGCGCKSTSVOLLRQILYDPVQGVLFDDYDAKELVNOMLRSGOIALYPPQEPVLFN 1042
Cc Db 1061 TLALVGSSGCGCKSTVOLLREYDPMASVFLDEGKEIRLOLVNOMRQALDGLVYSOEPIFLD 1120
Cc Qy 1043 CSAIENINYGDNSRVVPIDEITKEANMANINHSFIEIGLPEKYNTOVGLGAOLSGGQXORL 1102
Cc Db 1121 CSAIENINAYGDNRSRVYSIEEYIRAKKENIHOFIDSLDPKINTFRGDXGTOLSGGQCKRT 1180
Cc Qy 1103 AIALRALQPKILLDEATSAIUNDSSEKVVYOHALDKARTGTCVLTVTRLSAIONADVLV 1162
Cc Db 1181 AIALRALVQPHILLDEATSAIDTSEKRYVGEALDKAREGRFCVIAHRLSTIQNADLIV 1240
Cc Qy 1163 VLNHGKIKEGQGHQELNKRNDYEFVLVAQS 1193
Cc Db 1241 VIONGKVEKHGTHQOLLQAKGIFYSMVSVOA 1271

Cc RESULT 6
Cc ID MDR3_HUMAN STANDARD; PRT; 1279 AA.
Cc AC P21439;
Cc DT 01-MAY-1991 (Rel. 18, Created)
Cc DT 01-MAY-1991 (Rel. 18, Last sequence update)
Cc DT 30-MAY-2000 (Rel. 39, Last annotation update)
Cc DE MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3).
Cc GN ABCB4 OR PGY3 OR MDR3.
Cc OS Homo sapiens (Human).
Cc OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Cc MM Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Cc NX NCBI_TaxId=9606;
Cc RN [1]
Cc RP SEQUENCE FROM N.A.
Cc RX MEDLINE=89138016; PubMed=2906314;
Cc RA van der Bliek A.M., Koolman P.M., Schneider C., Borst P.;
Cc RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
Cc RL Gene 71:401-411(1988).
Cc RN [2]
Cc RP SEQUENCE OF 856-1279 FROM N.A.
Cc RX MEDLINE=8811519; PubMed=289268;
Cc RA van der Bliek A.M., Baas F., ten Houte de Lange T., Koolman P.M.,
Cc RT van der Velde-Koerts T., Borst P.;
Cc RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
Cc RT gives rise to alternatively spliced mRNAs in liver.";
Cc RL EMBO J. 6:3325-3331(1987).
Cc RN [3]
Cc RP GENE STRUCTURE.
Cc RX MEDLINE=91161629; PubMed=2002063;
Cc RA Lincke C.R., Smit T.J.M., van der Velde-Koerts T., Borst P.;
Cc RT "Structure of the human MDR3 gene and physical mapping of the human
Cc RT MDR locus.";
Cc RL J. Biol. Chem. 266:5303-5310(1991).
Cc CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
Cc CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT
Cc CC CAPABLE OF CONFERRING DRUG RESISTANCE. MEDIATES THE TRANSLLOCATION
Cc CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
Cc CC HEPATOCYTE.
Cc CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
Cc CC -1- DISEASE: DEFECTS IN PGY3 ARE A CAUSE OF PROGRESSIVE FAMILIAL
Cc CC INTRAHEPATIC CHOLESTASIS TYPE IIT (PFIC), A FORM OF AUTOSOMAL
Cc CC RECESSIVE LIVER DISORDERS. CHARACTERIZED BY EARLY ONSET OF
Cc CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
Cc CC ADULTHOOD.
Cc CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
Cc CC (ABC TRANSPORTERS). MDR SUBFAMILY.
Cc CC -----
Cc CC This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
Cc CC the European Bioinformatics Institute. There are no restrictions on its
Cc CC use by non-profit institutions as long as its content is in no way
Cc CC modified and this statement is not removed. Usage by and for commercial
Cc CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC or send an email to license@slb.ch).
CC
DR EMBL, M23234; AAA36207.1; -.
DR EMBL, X06181; CAA29547.1; -.
DR PIR, J50051; DVH03.
DR PIR, A42213; A42213.
DR HSSP; P13569; 1MBD.
DR MIM, 121060; -.
DR MIM, 602347; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001140; ABC_transportr_tmem.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_Ltran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multisite family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 354 POTENTIAL.
FT TRANSMEM 355 711 POTENTIAL.
FT DOMAIN 712 732 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.
FT TRANSMEM 854 873 POTENTIAL.
FT TRANSMEM 937 956 POTENTIAL.
FT TRANSMEM 976 993 POTENTIAL.
FT DOMAIN 994 1279 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (BY SIMILARITY).
FT NP_BIND 1069 1076 ATP (BY SIMILARITY).
FT REPEAT 1 640 ATP (BY SIMILARITY).
FT REPEAT 641 1279
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1093 1093 V -> VEFDFGFO (IN REF. 2).
SQ SEQUENCE 1279 AA; 140682 MW; 3D56C98B5C8D6087 CRC64;

Query Match 58.3%; Score 3504.5; DB 1; Length 1279;
Best Local Similarity 55.2%; Pred. No. 1.6e-196;
Matches 676; Conservative 235; Mismatches 276; Indels 37; Gaps 8;

QY 1 MILGLASLVNGACPLPMLPVLYGEMSDNLISCLVQNTYSF---FRLLT----- 47
   | | | | | : : : : : | | | | | : : | | | : | |
DB 57 MSLGTMIAIAGSGGJPLMMIVFGEMLDKFVD---TAGNFSPPVNFSLSLNPKRILEEE 112
   | | | | | : : : : : | | | | | : : | | | : | |
QY 48 -----YVVGIGVALILGYIGIISMLITTAAROTRRIRKRFHSVLAODIGFDSODICE 101
   | | | | | : : : : : | | | | | : : | | | : | |
DB 113 MTRVAYYISGLGAGVLAAYIOVSFTWTLAAGQIKRKIRKFFHALLROETIGFNDINDTIE 172
   | | | | | : : : : : | | | | | : : | | | : | |
QY 102 LNTTRMT-DIDKISDIDGKIALLPONMSTFFSIGLAVGLVKGKLLTVLTSPILMASNA 160
   | | | | | | | | | | | : : : : | | | | | : : | | | : | |
DB 173 LNTRLTDDISIKTSIEIGDKVGMFPAVAFPGFIVGIRGKLLTVLMAISPILGLSNA 232
   | | | | | | | | | | | : : : : | | | | | : : | | | : | |
QY 161 ACSRWVLSLTSKELSAYSKAGAVAEVLSIRTVTAFRQAEKELORYTONLKDAKDFGIK 220
   | | | | | : : : : | | | | | | | | | | | : : : : | | | | |
DB 223 VMAKITLSAFSDKELAAVYAKAGAVAEALGALRTVTAFGQNGLEBRYOKHLENAKEIGIK 292
   | | | | | : : : : | | | | | | | | | | | : : : : | | | | |
QY 221 RTIASVSLGAVYFFPMNGNTGGLAEFYRGISLLNGEPGTYIGTVLAVFSPVSHSSYICIGNA 280
   | | | | | : : : : | | | | | | | | | | | : : : : | | | | |
DB 293 KAISANISMGIAFLILYASYALAFYGSTLYTSKE--YTIAGMAYVFESILLGASVSGOA 350
   | | | | | : : : : | | | | | | | | | | | : : : : | | | | |
QY 281 VPHFTFATLGAAGAHIFIOVDIKKESINSEFAGYKPEISIECTVEPKANFSVPSPRSIK 340
   | | | | | | | | | | | : : : : | | | | | : : | | | : | |
DB 351 APCIDAFANMARGAAYVIFDIIDNNRKIDSFSRGKRPDSIKKNLEFNDFVHESTPSANRYK 410
   | | | | | | | | | | | : : : : | | | | | : : | | | : | |
QY 341 ILKGLINLKSGETVALVGLNGSGKSTVQALLQRLYDPDDGFIWVDENDIRALNVRHYD 400

```

```

Db 411 ILKGNLKVSGQYVALVSSGCCSTVQLQRLYDDEGINIDGDIRFNFNNTLRE 470
Oy 401 HIGVVSQEPVLFGTTISNNIKYGRDVTDEMERARARANAYDIMEPPNKNFTLVGEGK 460
Db 471 IIGVVSQEPVLFGTTISNNIKYGRDVTDEMERARARANAYDIMEPPNKNFTLVGEGK 530
Oy 461 AOMSGGOKRIATATARNRKLILDEATSLDSEKSNVQAALERSKGTITVVAHR 520
Db 511 AOMSGGOKRIATATARNRKLILDEATSLDSEKSNVQAALERSKGTITVVAHR 590
Oy 521 LSTIRADLIYTKDKLAEGAAHAEKRLQYSLV-----NSQDKKKEQME 571
Db 591 LSTIRADLIYTKDKLAEGAAHAEKRLQYSLV-----NSQDKKKEQME 650
Oy 572 SMTYSTERKTSNLPVSHYSKIK-SDFIDKA-EESTQSKREISLPEVSLKTIKLNKPEMF 629
Db 651 TRMAPNGKSRLEFRHSTQKNLNSQMCOKSLDVEFDGLEAVNPVPSFLKVLKNTKTEMPY 710
Oy 630 VILGFLASVLNGTVHPVPSIIFAKITTFGNNDKTLTKHDAEYEMIVYLGVICFVSF 689
Db 711 FVVGTVCAIANGGLQPARSVLSEITLIFPGDDAVKQCKNIFSLIFLGLIISFFTFE 770
Oy 690 MOGLFYGAGELITWRLRHLAFKAMLYODIAFDEKENSSTGGLTTLAIDIOIGATGS 749
Db 771 LOGFLFYGAGELITWRLRHLAFKAMLYODIAFDEKENSSTGGLTTLAIDIOIGATGS 830
Oy 750 RIGVLTQATNMGSLVSIISFYGMETFLISIAVLAVTGMETAAMTGFANKDKOELK 809
Db 831 RLATIAQNIANGTGIISFYGMETFLISIAVLAVTGMETAAMTGFANKDKOELK 890
Oy 810 HAGKATATALENIRITVSLTRKAFQOMYEMLOTOHRTSKKAOIGSCVAFSAFTF 869
Db 891 HAGKATATALENIRITVSLTRKAFQOMYEMLOTOHRTSKKAOIGSCVAFSAFTF 950
Oy 870 AYAGFRGAYLIQAGRMTPCEGMFVFTFAIAYAGNAIGTLVLAIEVSKAKGAHLFLAL 929
Db 951 AYAGFRGAYLIQAGRMTPCEGMFVFTFAIAYAGNAIGTLVLAIEVSKAKGAHLFLAL 1010
Oy 930 LEKRPIDRSOEGKKPPTCEGNLEFREVSEFFYPCRPDVFILRGLSLSTIERKTVAFVGS 989
Db 1011 FEROPILDISYSEGLKPKFEGNITFNEVFNYPTRANVPVLOGLSLEVKKQOTLALVS 1070
Oy 990 SGCGKSTSVOLLQRLYDVOGVLFQDQAKLNVQWLRSAIATYPOEVVLNCSIAEMI 1049
Db 1071 SGCGKSTSVOLLQRLYDVOGVLFQDQAKLNVQWLRSAIATYPOEVVLNCSIAEMI 1130
Oy 1050 AYGDNRVVPVLDIEKEAANAANHSIFEGLEPERYTOVGLKGAQLSGGOKRIATARALL 1109
Db 1131 AYGDNRVVPVLDIEKEAANAANHSIFEGLEPERYTOVGLKGAQLSGGOKRIATARALL 1190
Oy 1110 OKPKILLIDEATSLDSEKSNVQAALERSKGTITVVAHR 1169
Db 1191 ROPKILLIDEATSLDSEKSNVQAALERSKGTITVVAHR 1250
Oy 1170 KEQGHQELLRNDYFKLVNQS 1193
Db 1251 KEQGHQELLRNDYFKLVNQS 1274

```

```

RESULT 7
MDR2_MOUSE
ID MDR2_MOUSE STANDARD: PRT: 1276 AA.
AC P21440:
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).
GN PG2 OR PG2-2 OR MDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-88302195; PubMed-3405218;
RA Gros P., Raymond M., Bell J., Housman D.;
RT Cloning and characterization of a second member of the mouse mdr
RL Mol. Cell. Biol. 8:2770-2778(1988).
RP [2]
RP SEQUENCE OF 1-23 FROM N.A.
RC STRAIN-BALB/C;
RA Kirschner L.S., Horwitz S.B.;
RL Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS
CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03398; AAA9516.1; -
DR EMBL: M74151; AAA39515.1; -
DR PIR: A30409; DVMS2.
DR HSSP: P13569; INED.
DR MGD: MGI:97569; Pgy2.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 75 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 753 773 POTENTIAL.
FT TRANSMEM 829 849 POTENTIAL.
FT TRANSMEM 850 870 POTENTIAL.
FT TRANSMEM 933 953 POTENTIAL.
FT TRANSMEM 970 990 POTENTIAL.
FT DOMAIN 991 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (BY SIMILARITY).
FT NP_BIND 1066 1073 ATP (BY SIMILARITY).
FT REPEAT 1 635
FT REPEAT 636 1276
SQ SEQUENCE 1276 AA; 140332 MW; A6C38DCD2C118EEF CRC64;

```

```

Query Match 57.9%; Score 3480.5; DB 1; Length 1276;
Best Local Similarity 54.7%; Pred. No. 3.9e-195;
Matches 671; Conservative 236; Mismatches 278; Indels 41; Gaps 6;

```

```

Oy 1 MIIIGILASLVNACUPLMPLVIGENSNDLISGLVQNTYVSF-----FRILT----- 47
Db 54 MFLGLTALTAHNSGLPLMIVIGEMTKRFVD-----NFGNLSLPVNFSLNLPRIILE 107

```

Qy	48	- - - - -	YVGVGVAALLFGTIOJSTWITTAARQTRIKROFPHSVLADIDMFPSODI	99
Db	108	EMTRRAYAYISGGVVAAYIOWSFMTLAAAGRIKRIKOFKHFALLIKROEMGFIDIKT	167	
Qy	100	GELNTRMT -DIDKISGIGDKIALLFQNMSPFISLAVGLKQKMLTYLSTSPILMAS	158	
Db	168	TELNTRLTDDVKSISTSGIDGKGMFFQALATTFAGFIVGFRKGLTYLTMALSPILGS	227	
Qy	159	AAACSRWVISTLSTKESLAYSKAGAAEEVLSIRTVAFRAQKELORYTQNLKADKFG	218	
Db	228	TAYMAKILTSFSDKELAAAYAKAGAAVEAPGARITVIAFGQONKELEBYOKHLENAKKIG	287	
Qy	219	IKRTIASKVSGLAVYFFPMNGTGLAFWYOTSLINGEGYITIGYLAFFSVIHSSTCIG	278	
Db	288	IKKASINIMSGIAFLFLIAYASVALAFWYSTLVISKE -YTIGNAMTYFFSILIGASVSG	345	
Qy	279	AAVPHETFLAIGAAPFHFOVODKPSIDNSTFAGYKPESTIEGVEFKNVSFNPYRPS	338	
Db	346	QAAECPIAENAGAAVLFEDIIIDNNPKIDSEKGNHFPDINKGLEPSDVHSTYPRAN	405	
Qy	339	IKILKLNLRKISGETVAVLVLGNGSGSKSTVOLLRLYDPDPDGITMDENDIRALNRY	398	
Db	406	IKILKGLNLTAKVSGGVYALVNGSGCKSTVOLLRLYDPPEGRKISIDGDIRNFENRCL	465	
Qy	399	RDHIGVSOEPUVLFGTTISNNIKYCGDDYTDDEMRAPAREANADYFTEFNPKNNTLYGE	458	
Db	466	REIIGVISOEPUVLFSTTIENIRYTGNGYTMJEIKAYEANAADYFIKKLEOKDKDLYGD	525	
Qy	459	KGAMSGGOKORIALAVALRNPKILILDEATSLDSESKSVYOALFEKASKGRTTYVA	518	
Db	526	RGAOLSGGOKORIALAVALRNPKILILDEATSLDSESAVVOALDKAKEGRTTYIA	585	
Qy	519	HLRSTISADLIYVLKQDMLEAKGHAHELMARKGLYLSYSQO-----DIRKAD	567	
Db	586	HLRSTINADVIAFGEDGVIVEGSHSELMKKEGTYFRLYNMQTGSQLSEEFVELSD	645	
Qy	568	EOMESMYSYTERKTNLSLPLHSVAKSISDPTIDKAESTOSKESLSEPVSLKILKLKNPEW	627	
Db	646	EKAAGDVAAPMGMKARIFRNSTRKSLKSPQNRDELETENDJANPPVPSFLVVLKNTKEW	705	
Qy	628	PEVVLGTLASVLTNGTVHPVPSIIFAKIIMFQGNKDTIKHDAELYSIMFVLIVGICVS	687	
Db	706	PEVVGVCALANALOAPASIIILSEMLIFGGDDAVAKQKCNFSLVPIGLVGLSFT	765	
Qy	688	YFMQGLEFGRAGEIITLWRLHLAFKMLIYODIAMPDEKENSGLTTIILADIADIOGAT	747	
Db	766	FFLOGFTRPKRAGEIITLRLTSLMAFKMLKODMSHPDKHNSGALSTRLADDAOYGAT	825	
Qy	748	GSRIQVLTQNTNMGSLVISIIFYGEMTFLLISAPVLAVTGMETLAAMTGFAANKOE	807	
Db	826	GTKLALLANONTANGTIIISIFYGQOLTFLLISVPVLAAGVIEKMLGNKAKRDKKE	885	
Qy	808	LKHAGKATATELENIIRTVSLTREKAFEQMYEEMLOTORHNTSKRAQIIGSOYAFSAFI	867	
Db	886	MEAGKATATELENIIRTVSLTOERKFESYVEKLEGPARNSVKRAHHIYGTFSTISOQFM	945	
Qy	868	YFAAYAGRFCAZYIQAQRYMTPEGMEYFPAIYAGMAIIGKLYLVAPEXSKASGAHLF	927	
Db	946	YFSAVAGCRBSSTYLVNGHHRKFDVILVEFAYLVGVALGHAHSSAPADPAKAKSAATLF	1005	
Qy	928	ALLERKPNIDRSOEGKKPDTCBGNTLEFREVSFFYPCRPDVEFLNGLSLSTIERGTVAVF	987	
Db	1006	SLEFROPLIDYSSEGLMPDKFEGSVTFNEVYFNPTRANVPVLQGLSLEVKKGGTIALV	1065	
Qy	988	GSSGGKSTYVOLLRLYDPYQGVLFDDYDAKELAVQNLBEOIALVPOEPUVLFNCSIAE	1047	
Db	1066	GSSGGKSTYVOLLERYDPAAGSVLLDDGEAKLWQMLRAQOLGVESEPLLEFCSIAE	1125	
Qy	1048	NAYCDNSRVPLDEIEEANAANIHSEFJIGLEPEKYNTOVGAKLOSGGOKORIALARA	1107	
Db	1126	NAYCDNSRVPLDEIEBRAKEANIHIEFTLQOKNTRYKDGKTYOLSGGOKORIALARA	1165	

Oy	1108	LLOKPKILLIDEEASALDNSEKVVQAHLDKAPRGRCGLVTHRLSAQNMDLIVYLHG	1167
Dd	1186	LIRDPKVLLEDAFVSALTDESKESVYQEAFLDAREGRFCIVIAHKLSTIQNADLIVIEHG	1245
Oy	1168	KIKEGCTHOELLRNNDIYFKLVNAOS	1193
Dd	1246	KYKEHGTNQQLLAKQIGITFSVMNIQA	1271
RESULT	8		
ID	MDR1_RAT	STANDARD:	PRT: 1277 AA.
AC	P43245;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).		
CN	ABCB1 OR PGY1 OR MDR1 OR MDR1B.		
OS	Rattus norvegicus (rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
NCBI	_taxid=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92039081; PubMed=1682220;		
RA	Silverman J.A., Kaunilo H., Gant T.W., Thorgerirsson S.S.;		
RT	"Cloning and characterization of a member of the rat multidrug		
RT	resistance (mdr) gene family.";		
RL	Gene 106:229-236(1991).		
CC	-! FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED		
CC	DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.		
CC	-! SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.		
CC	-! PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE		
CC	LINKER DOMAIN.		
CC	-! MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE		
CC	RELATED BUT DISTINCT CELLULAR GENES.		
CC	-! SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY		
CC	(ABC TRANSPORTERS). MDR SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; M81855; -; NOT ANNOTATED_CDS.		
DR	InterPro; IPR003593; AAA.		
DR	InterPro; IPR003439; ABC_transporter.		
DR	InterPro; IPR001140; ABC_transprt_tmnm.		
DR	InterPro; IPR001687; ATP_GTP_A.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER_2.		
KW	ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;		
KW	Multigene family.		
FT	DOMAIN	1	47
FT	TRANSMEM	48	71
FT	TRANSMEM	119	139
FT	TRANSMEM	199	219
FT	TRANSMEM	296	316
FT	TRANSMEM	327	347
FT	DOMAIN	348	709
FT	TRANSMEM	710	730
FT	TRANSMEM	755	775
FT	TRANSMEM	832	852
FT	TRANSMEM	854	874
FT	TRANSMEM	937	957
FT	TRANSMEM	968	988
FT	DOMAIN	989	1277
FT	NP_BIND	426	433
FT			

```

FT NP_BIND 1070 1077 ATP (POTENTIAL).
FT REPEAT 1 635
FT CARBOHYD 636 1277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 1277 AA; 141386 MW; 8AFDD619D2934C1 CRC64;

Query Match 57.8%; Score 3476.5; DB 1; Length 1277;
Best Local Similarity 35.4%; Pred. No. 6,6e-195;
Matches 681; Conservative 233; Mismatches 272; Indels 43; Gaps 11;

OY 1 MIIIGLILVNGACPLPLPLVGEKSDNLI-----SGCIVQNTNYSFRL-- 45
DB 49 MALGLTAAIIHGTLLPLMLPLVFGYMTDSPTSPRDPHSRAITNOSSEINSTHVSOTSLFE 108
OY 46 -----TLVYVIGVAAALIFGYIQLSWITTAARQTKRIKOPHSVLAODIGWPDSCDIG 100
DB 109 DMAMVAYYYTGIGAGVLLVAYIQVSLWCLAGRQIHKTIRQKFHAIMNGEIMWPDVNDAG 168
OY 101 ELNTRMT-DIDKISDGIQDKTALLFQNMSTSIGLAVLGWKLTLVTLSTPLIMASA 159
DB 169 ELNTRITDVSKRINDGIDKLGMPFOSITTFSGFIIIGISGKMLTLVLAVLSPILGLSS 228
OY 160 AACSRVYLSLSKELSAYSKAGVAEVLSSIRVYIAFPAQOEKELORTONLKDADKDEGI 219
DB 229 AMMAVAVLSFINKELQATKAGAAVEVLAALRTVIAFGQKLEKRYKNLEAKRVGI 288
OY 220 KRTIASKSLGAVVFPFMNGTGLAFWGTSLILNGEPYTGTVLAVFSPVSHSSYCGA 279
DB 289 KKAITANISIGIAYLLVAVSALAFWYGSIVLSNE--YSIQVTLVFPSSILGFFSIGH 346
OY 280 AVPHETAIARGAAPHIPIQVTDKPSIDNSTAGIKPEPSIEGYVEFKVNSNYSPRSPI 339
DB 347 LAPNEAFAANAGAAEYEFKIIDNEPSIDSFSTKGHKPSINGNEFEKVVYFNPSRSYV 406
OY 340 KILKGLNIRIKGETVALVGLNGSKSTVYVOLLORLYDDDFIWDENDIDALVNRHR 399
DB 407 KILKGLNIRIKVSGGVALVGLNGSKSTVYVOLLORLYDDDFIWDENDIDALVNRHR 466
OY 400 DHIGVSOEPLVLEGTTSINNIKYGRDVTDEEMERAREANAYDFIMEPNKFNLTLYGK 459
DB 467 ELIGVSOEPLVLEGTTSINNIKYGRDVTDEEMERAREANAYDFIMEPNKFNLTLYGK 526
OY 460 GAOMSGGOKORAIARALVRNPKIILDEATISALDESASVQAALAKASCKRTTIVAH 519
DB 527 GAOMSGGOKORAIARALVRNPKIILDEATISALDESASVQAALAKASCKRTTIVAH 586
OY 520 RLSTIRADLIYTLKDMGMAEKGAHMAELAKRGVLYSLVNSQDIDKADROMESMTYSTER 579
DB 587 RLSTIRADLIYTLKDMGMAEKGAHMAELAKRGVLYSLVNSQDIDKADROMESMTYSTER 645
OY 580 KTNLSPLHNSVKSISDFIDKA-----EESTOSKE--ISLPEVSLKILKLNPK 625
DB 646 DTGASELSEES-KSPILRISIRSIHRQDDEBRLLSKSEDEDEVPMVSPFOILLKNTS 704
OY 636 EMPFVVLGTSLAVLNGVYHPVSIIFAKITIMFG--NNDKTTIKHDAEIVSMITFVLGVC 684
DB 705 EMPFVVLGTSLAVLNGVYHPVSIIFAKITIMFG--NNDKTTIKHDAEIVSMITFVLGVC 764
OY 665 FVSYMOGLFVGRAGEIILMLRLHAFKAMLYODIAMPEKNSNGGLTIIAIDAQIO 744
DB 765 FVSYMOGLFVGRAGEIILMLRLHAFKAMLYODIAMPEKNSNGGLTIIAIDAQIO 824
OY 745 GATGSRIGVLTQONATNMGISLVITS--FIYGEWTEFLLISIAVLAVTGMIEAAMTFAN 802
DB 825 GAMGSRIAVYVONVANLNGIILSLVLVYGMQTLTLLVLIIVLIVGIIIEKMLSGQAL 884
OY 803 KKOELKNAKATATALENIRIVSLTRKAFEQMTEEMKQTOHNTSKKAOITGSCYAF 862
DB 885 KKKKELEISGKTATATALENIRIVSLTRKAFEQMTEEMKQTOHNTSKKAOITGSCYAF 944
OY 863 SHAFIYFAAAGFRGAVLIQAGRMTPBGMFLVFTALAVGAMALGKTLVLAREYSKAKSG 922

```

```

DB 945 TQAMIFYSAACFRGAVLYARELMTFEENVMLVFSAYVFGAANAAGNTSSFADYAKAVS 1004
OY 933 AAHLFALLEKKNIDRSROEKKPDTCEGNLEFREVSEFFPCRPDVFILRGISLSIERCK 982
DB 1005 ASHIIIGIIEKPEIDISYSTEGLKPNMLEGNKFNKFPYPRPPIVYLOGISFEVKKQ 1064
OY 983 TPAFVSSCGCKSTSVOLLORLYDDPVOGVLPDGVADKELANQWMLRSQAIYPOEPVLEN 1042
DB 1065 TRLVGLSSCGCKSTVOLLORLYDDPVOGVLPDGVADKELANQWMLRSQAIYPOEPVLEN 1123
OY 1043 CSAENIAYGDNRSVYPLDEIKEAANAANHSFIEGLPEKNTVOYGLKGAQLSGGOKRL 1102
DB 1124 CSAENIAYGDNRSVYPLDEIKEAANAANHSFIEGLPEKNTVOYGLKGAQLSGGOKRL 1183
OY 1103 AIALRALKPKILLDEATSAIDNDESEKVVQALDKARTGRCVLVYTHLSAIONADLIY 1162
DB 1184 AIALRALKPKILLDEATSAIDNDESEKVVQALDKARTGRCVLVYTHLSAIONADLIY 1243
OY 1163 VJHNGIKRQGTQHOELLRRDRIYFKLVNA 1191
DB 1244 VJHNGIKRQGTQHOELLRRDRIYFKLVNA 1272

RESULT 9
MDR3_CRIGR STANDARD; PRT: 1281 AA.
AC P23174;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3).
GN PG3 OR PG3.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_Taxid=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Saragol F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
   gene family."
RL DNA Seq. 2:89-101(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC -1- DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- MISCELLANEOUS: PG3 ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PG1 AND PG2 CAN MEDIATE MDR, WHILE PG3 APPARENTLY
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL: M60042; AAA68885.1; -.
CC HSPD: P13569; INBD.
CC InterPro: IPR003593; AAA.
CC InterPro: IPR003439; ABC_transporter.
CC InterPro: IPR001140; ABC_transporter_tmem.
CC InterPro: IPR001687; ATP_GTP_A.
CC Pfam: PF00664; ABC_membrane_2.
CC Pfam: PF00005; ABC_tran; 2.
CC SMART: SM00382; AAA; 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;

```







DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001140; ABC\_transportr\_tmemb.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR SMART: SM00382; ABC\_tran; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
Multigene family.  
KW DOMAIN 1 54 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 119 139 POTENTIAL.  
FT TRANSMEM 188 208 POTENTIAL.  
FT TRANSMEM 215 235 POTENTIAL.  
FT TRANSMEM 296 316 POTENTIAL.  
FT TRANSMEM 325 345 POTENTIAL.  
FT DOMAIN 346 709 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 710 730 POTENTIAL.  
FT TRANSMEM 755 775 POTENTIAL.  
FT TRANSMEM 831 851 POTENTIAL.  
FT TRANSMEM 852 872 POTENTIAL.  
FT TRANSMEM 935 955 POTENTIAL.  
FT TRANSMEM 972 992 POTENTIAL.  
FT DOMAIN 993 1278 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 426 433 ATP (POTENTIAL).  
FT NP\_BIND 1068 1075 ATP (POTENTIAL).  
SQ SEQUENCE 1278 AA; 140655 MW; DAFB6BE745AF73BF CRC64;

Query Match 57.0%; Score 3428.5; DB 1; Length 1278;  
Best Local Similarity 54.0%; Pred. No. 4.2e-192;  
Matches 663; Conservative 240; Mismatches 282; Indels 43; Gaps 7;

QY 1 MILGLIASLVNGACPLPMLVLGEMSDNLISGLVOTNTYSE-----FRLLT----- 47  
DB 54 MILGTMAIAHSGPLMHVGEEMDKFYD-----NAGNFSLPVNFSLMNLNPKRIIE 107  
QY 48 -----YVGIGVALIFGYIQISLMTITTAARQTRKRIKQFFHSVLADIDCFSCDI 99  
DB 108 EEMTRVAYVYSGIGCVLLAAYIQVSEFWTLAAGROIRIKIQKFFHAILRQEMGFIDIKGT 167  
QY 100 GELNTEMT-DIKISDIDGDKIALLFONNSTFISGLAVGLVGMKTLTVTLSTPLMAS 158  
DB 168 TELNTEPLTDIKISDIDGDKIALLFONNSTFISGLAVGLVGMKTLTVTLSTPLMAS 227  
QY 159 AAASRMVLSLSKELSAVSKAGAVAEVLSIRTVIAFRAQEKELQRYQNLKDAKDFG 218  
DB 228 TAVMAKILSTFSDKELAAAKAGAVAEALGATRIYIARGGQNKELERYQKHLNKKIG 287  
QY 219 IKRTIAKSVSLGAVYFPMGTGGLAFWGTSLILNGEPYTGTVLAVFSSVHSYCIQ 278  
DB 288 IKKAIISANISMGIATFLIYASVALAFWYSTLVISKE--YTIGNAMTVFSSILIAFVSQ 345  
QY 279 AAVPHETFAIRGAFFHFOVTDKRPSTDNFTAGYKPESTIGYEFKFNVSNNYSRPS 338  
DB 346 QAAPCTIDAPPNAGAVYFIDIIIDNPKIDFSEKGHKPDISIGNLEFSDVHSYPSRAN 405  
QY 339 IKILKGLNIRIKSGETVALVGLNGSGSPYVOLLQRLYDDDDGFIMVEDIDIALVNRHY 398  
DB 406 IKILKGLNIRIKSGETVALVGLNGSGSPYVOLLQRLYDDDDGFIMVEDIDIALVNRHY 465  
QY 399 RDHIGVASEPVLGTTTISNNIKYGRDVTDEEMERAAERANAYDFIMEFPNKNFTLVGE 458  
DB 466 REFIGVASEPVLGTTTISNNIKYGRDVTDEEMERAAERANAYDFIMEFPNKNFTLVGE 525  
QY 459 KGAWSGGOKORAIARALVLRNPKIILDEATSAIDSEKSAVAQALAEKRSKRTIYVA 518  
DB 526 RGAQISGGOKORAIARALVLRNPKIILDEATSAIDSEKSAVAQALAEKRSKRTIYVA 585  
QY 519 HRLSTIRSDALVTYLDGMLAEKGAHAELMAKGLVYLSWQ-----DIKKAD 567  
DB 586 HRLSTIRSDALVTYLDGMLAEKGAHAELMAKGLVYLSWQ-----DIKKAD 645

QY 568 EQMESMTYSEKRTNSLPLHSVKISKD--FIDKAEOSTQSKETISUPEVSLIKTLKLNK 625  
DB 646 EKAGGAVAPNGMKARLFERNSTKSLKSSRAHQNRDLVETNELDANPVPVSLVLYLNT 705  
QY 626 EMPEVVLGTSLASVLTNGTVHPVPSIIFAKITTMGNDKTTLKDAEYISMIFVILGICE 685  
DB 706 EMPEVVLGTSLASVLTNGTVHPVPSIIFAKITTMGNDKTTLKDAEYISMIFVILGICE 765  
QY 686 VSYEMOGLFEGRAGEILYMLRHLAFKAMLYODIANFDEKENSTGGITTLAIDIAIQG 745  
DB 766 FTFPLGTFEGKAGEITLTLRLSMFAKAMLRDMSWFDHKNSTGALSTFLATDAAOVQ 825  
QY 746 AGSRIGVLTQNTNMNGSLVITSTIGWEMTEFLISAPLAATGMETFAAMGPNKDK 805  
DB 826 AGTRIALIAQNTANLGTGLISIFIGWDTLLLSVPPPIAVAGIVEKMLGNKRDK 885  
QY 806 QELKHAGKATATEALNTRTVSLTRKAEFQMEEMIQTRHRTSKKAOITGSCVAFSHA 865  
DB 886 KEHEAAGKATATEALNTRTVSLTRKAEFQMEEMIQTRHRTSKKAOITGSCVAFSHA 945  
QY 866 FIYFAVAGRFAGVLIQAGMTPEGMFIVETAIYAGAMAIGTVALAPYSAKSGAAH 925  
DB 946 FMVFSYAGCFRFGSYLVNNGHMRKDYILVFSALVAGVALGHASSFAPDYAKAKLSAAV 1005  
QY 926 LFLALEKRPIDRSQDEGKKPDCEGNLEPREVSFFPCPDYFIIGLSTIERKTYA 985  
DB 1006 LFLALEKRPIDRSQDEGKKPDCEGNLEPREVSFFPCPDYFIIGLSTIERKTYA 1065  
QY 986 FVSSGCGKSTSVOLLQRLYDPQOVLPFDGVAKELNVMQLRSQALYAOEVEFLNCST 1045  
DB 1066 LVSSSGCGKSTSVOLLQRLYDPQOVLPFDGVAKELNVMQLRSQALYAOEVEFLNCST 1125  
QY 1046 AENIAYGDNRRVPLDEIKKANANANISFTIEGLPEKYNTOVGLKGAQISGGOKORAI 1105  
DB 1126 AKNIAYGDNRRVPLDEIKKANANANISFTIEGLPEKYNTOVGLKGAQISGGOKORAI 1185  
QY 1106 RALLQPKILLDEATSAIDSEKVAQALDARGGRCLVYTHRLSATONADLYVLH 1165  
DB 1186 RALLQPKILLDEATSAIDSEKVAQALDARGGRCLVYTHRLSATONADLYVLH 1245  
QY 1166 NGKIKEGTQHOELLRNNDYFKLVNAOS 1193  
DB 1246 NGKIKEGTQHOELLRNNDYFKLVNAOS 1273

RESULT 11  
AB11\_HUMAN STANDARD; PRT; 1321 AA.  
AC 095342; Q90NB2.  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11).  
GN ABCB11 OR BSEP.  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OC NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT PFIC2 GLY-297.  
RX MEDLINE-99021377; PubMed-9806540;  
RA Strautnieks S.S., Bull L.N., Knisely A.S., Kocoshis S.A., Dahl N.,  
RA Arneil H., Sokal E., Dahan R., Childs S., Ling V., Tanner M.S.,  
RA Kagalwalla A.F., Nemeth A., Pawlowaka J., Baker A., Miel-Vergran G.,  
RA Fagerman N.B., Gardner R.M., Thompson R.J.;  
RT "A gene encoding a liver-specific ABC transporter is mutated in  
RT progressive familial intrahepatic cholestasis".  
RL Nat. Genet. 20:233-238(1998).  
RP SEQUENCE FROM N.A.  
RA Mol O., Hootveld G.J.E.J., Jansen P.L.M., Muller M.;  
RT "Cellular localization and functional characterization of the human



Db 1017 VERVISAVALSTALGAFSTYPSYAKAKISARFFQULLDRPPLISVYTAAGKMDNPFQ 1076  
QY 953 NIEFEVSEFFYCRPDVFLRLGLSIEKGTVAAYVSGGCKSKTSVOLLRLYDPVQO 1011  
Db 1077 KIDFDYCKFTYSPRDSQVLAISISPGOTLAFVSGGCKSKTSIQLLEFYDPDQK 1136  
QY 1012 VIFDGVDAKELVOMLRQIAIVPOEPVFNCSIAKENTAVGNSRVPYDEIKELANMAN 1071  
Db 1137 VNIHDHDSKVVAVQPLRSNIGVISOEPVLFACSIMDNIKYGNTEIPMERVIAAKQAO 1196  
QY 1072 IHSFTEGLEPEKNTQVGLKAOLSGGOKRQIAIARALLQPKILLDEATSAIDNSEKY 1131  
Db 1197 LHDFVMSLEPEKETVNGSGGSLRGEKORIAIARIVADPKILLDEATSAIDTSEKT 1256  
QY 1133 VOHALDKARTGRTCLAVYHRLSAIQNADLLVYLVHNGKIEGCTHOELLRNDRYELVNA 1191  
Db 1257 VOVALDKAREGTCTVIAHRLSTIQNADIIAVMAOGVIEKGTHERELMAKGYKLVYT 1316  
QY 1192 QS 1193  
Db 1317 GS 1318

RESULT 12  
AB11\_RAT STANDARD: PRT: 1321 AA.  
ID AB11\_RAT  
AC 070127;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)  
DE (SISTER OF P-GLYCOPROTEIN).  
GN ABCB11 OR BSEP OR SGP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
RX MEDLINE=9812048; PubMed=9545351;  
RA Gerloff T., Steiger B., Hagenbuch B., Madon J., Landmann L., Roth J.,  
RA Hofmann A.F., Meier P.J.;  
RT "The sister of P-glycoprotein represents the canalicular bile salt  
RT export pump of mammalian liver.";  
RL J. Biol. Chem. 273:10046-10050(1998).  
CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
CC INTO THE CANALICULUS OF HEPATOCYTES.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
CC SITU.  
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
CC ATP BINDING CASSETTE (ABC) DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: U69487; AAC40084.1; -;  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003439; ABC\_transport.  
DR InterPro: IPR001140; ABC\_transport\_tmem.  
DR InterPro: IPR001687; ATP\_GTP\_A.

DR Pfam: PF00664; ABC\_membrane: 2.  
DR Pfam: PF00005; ABC\_tran: 2.  
DR SMART: SM00382; AAA: 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER: 1.  
KW ATP-binding; Transmembrane; Transport.  
FT DOMAIN 1 62  
FT TRANSMEM 63 83  
FT TRANSMEM 84 147  
FT TRANSMEM 148 168  
FT TRANSMEM 169 215  
FT TRANSMEM 216 236  
FT TRANSMEM 237 240  
FT TRANSMEM 241 261  
FT TRANSMEM 262 310  
FT TRANSMEM 320 340  
FT TRANSMEM 341 353  
FT TRANSMEM 354 374  
FT TRANSMEM 375 755  
FT TRANSMEM 756 776  
FT TRANSMEM 777 794  
FT TRANSMEM 795 815  
FT TRANSMEM 816 869  
FT TRANSMEM 870 890  
FT TRANSMEM 891 911  
FT TRANSMEM 912 979  
FT TRANSMEM 980 1000  
FT TRANSMEM 1001 1011  
FT TRANSMEM 1012 1032  
FT TRANSMEM 1033 1321  
FT NP\_BIND 455 462  
FT NP\_BIND 1113 1120  
FT CARBOHYD 109 109  
FT CARBOHYD 116 116  
FT CARBOHYD 122 122  
FT CARBOHYD 125 125  
SQ SEQUENCE 1321 AA; 146257 MW; 5443F4E7B9FB1P6 CRC64;  
Query Match 47.2%; Score 2839; DB 1; Length 1321;  
Best Local Similarity 45.5%; Pred. No. 16-157;  
Matches 572; Conservative 249; Mismatches 363; Indels 74; Gaps 10;  
QY 1 MIIIGLILSYVNGACLPMLPVLYGEMSDNLI-----SGLVQNTNY----- 40  
Db 62 MIMGVCCALLHGMAQGIILFEGIMTDIFLKYDIROELEIPKACVNMNTVINSSPHQ 121  
QY 41 -----SFFRLTYVVGIGVAALIFGYIQLSLMTITTAARQTKRIKOFENS 85  
Db 122 NMNGVCGVGLVDESEMIFSGIYAGVMVLILGYFQIRLWYITGARQRRRKIYFR 181  
QY 86 VLAQDIGWFDSCDIGELNTRMT-DIDKISDGIDKIALLEFONNSTPSIGLAVGVKWK 144  
Db 182 IMMEIGWFECTSVGELNSRFADIEKINDAIDQLAHLFQRMSTAMCGILLFYGWK 241  
QY 145 TLVLTSLPLMASAACSMMVLSLSKELSAASKGAVAEVLSIRYIAARAQEK 204  
Db 242 TLVILAVSPILIGIAVIGLSIKTELEIKAAKAGSTADEVLSIRYVAAAGGK 301  
QY 205 QRTYQMLKAKDGIKRTIASKSVLSGAVYFFNMGITGLAFWYTSLLNCEPGYITGV 264  
Db 302 EREKRLVRAQRMGKWKGMVGFETGMCLIFCYALAFWYSTVLVD-EERYTGTV 360  
QY 265 AVFVSIIHSSYICGAAVPHFEFAIRGAFAHIFQVIDKKPSIDNSTAGYKRESETEGV 324  
Db 361 QIFLCVILAAAMNIGHASSCEIFSTGCSATNIFQTRDPVIDCMGSGYKIDRIKGEI 420  
QY 325 EEFNVSFNPSRPSIKILINGINLRKISGEVAVLVGNSGSKSVYVOLLRLYDPDGF 384  
Db 421 EFNVTFFHPSRDPVAILDLNSLVIRPGETTALVSSGAKSTALDILQRFDPGGMVT 480  
QY 385 VDENDIRALVRRHYRHIGVAVSEPVLFCTTISNNIKYGHADVTDDEMERARANAAYDF 444  
Db 481 LDQHDIRSLNIRLRQIIGVIEBPVLFSTTIENIRFGEDATMEDYVQAANDANAYNF 540



FT CONFLICT 633 633 T -> V (IN REF. 2).  
 SQ SEQUENCE 1321 AA: 146675 MW: 1555EBF175D32967 CRC64:

Query Match 47.1%; Score 2834; DB 1; Length 1321;  
 Best Local Similarity 45.8%; Pred. No. 2e-157;  
 Matches 576; Conservative 244; Mismatches 363; Indels 76; Gaps 12;

1 MILGILAVLNGACLPRLPLVLGEMSD----- 27  
 62 MEMSVCLLHGMAQPGMIIVFGILTFVEXDIEROELSPGKVCMMNTWIMNSFNO 121  
 28 NLISG--C-LVOTNTSPFRLTYVYGIVGVALINGYIOISLMTTAAQOTKIRKOPFH 84  
 122 NMTNSTSGCLVDINS-EVTKFSGIYAGVAVLLIGFQIRLWITVGAQIRKMKRYFR 180  
 85 SVLADIDMFPSCDIGELINTRT-DIDKISDIDGKIALLEPOMSTFSGIYAVLGKMK 143  
 181 RIMREIGMPDOTSVGELNSRSDINKIDEIADQMALFQIRLSTALSGLLGFRGK 240  
 144 LTVLTSLTPLMASAACSRMVISLTKELSAVSKAGAVAEVLSIRTVAFRAQKE 203  
 241 LTVLTAVSPLIGIAAIGLSVAFTLELKAAYAKGSIADDEVLSIRTVAAFGEMKE 300  
 204 LQRTONLKDADKDGIRKRTIASKVSLSGAVYFPMNGTYGLAFMYGSLILNGSPGTIGY 263  
 301 VEREKNLMEFQRMGIMGMWGFYGMWCLIFPCYALAEYGSRLVD-EGEYTPGT 359  
 264 LAVFVSIVSSYICGAAPHEFTFAIARGAAFIHQVIDKKPSINFSTGAKPESIEGT 323  
 360 IQIFLCVILIAAMNIGNASCIETFGSCAASSIFQITDRQVDMCMGSDYKLDRIKE 419  
 324 VEKRVNYSNRSRPIKTLKGLNLKISGEYVALVGLNSGKSYVOLQRLYDPDDGT 383  
 420 IEFNVTVEHYSPREVKILNMLSMVTKPGETAFVSSGAKSTALQLOFRFYDCEGMY 479  
 384 MVDENDIALNVRHRIHIGVSOEPLFGPTISNNIKYGRDVEDMEERAREANVD 443  
 480 TLGDHIDSLNRLRDIGIYEDPVLFTTIAENIRIGREATMEDIVQAKDNAMN 539  
 444 FIMEPNKFTLVGEGKAGMSGQKORIAIALAVRNPKILLIDETSALDESKSAYOA 503  
 540 FIMALPQGFDTLVGEGGMSGQKQVAIALALRKPKILLDMATSADNESEKAYOG 599  
 504 ALEKASKRITIVAHRLSTRSADLIYTLKDGMAEKGAHAELMAKRGLYSLVMSOD- 562  
 600 ALNKIOHGTITISVAHRLSTVRSADVIIIFGEHTAVERGTHELLEKRGVYMLVLSQ 659  
 563 -----IKKAD-----EOMESMTYSTERKNSLPLH-----SVKSI 592  
 660 EDNTHKEGIGKDTEDDTPERTFSRGSYODSLNASTRORKSQLSLSHEPPLAIGDH 719  
 593 KSDPIDKAESTQSKEISLPEVSLILKLNKPEMPFVLGLAVLGVIPVPSIIFA 652  
 720 KSSYEDRDNDVLEVEBPAPVR--RIIKYNISEMPYILVGLCAINCAVPIYISLFS 777  
 653 KIITWEGNNDKTTLKHADEIYSM--IYVILGVICFVSVMOGLEFGRAEIIITMLRLHIA 710  
 778 QILKFFSILVDE--QORSEIYSCLEFVILGVSLEFLOGLVSNPAKSGEILLTKRLKRG 835  
 711 FKAMLYODIAMPDEKENGSTGLTTLADIDAOIGATGSRIGVNLONMTNMLSYISRI 770  
 836 FKAMLRQDIDGMDLKNPVGVLTRLATDASOVGATSGQGMAMNSFTNITVAVALIATL 895  
 771 YGMENTPILSLIAPLAVTAMTGMTAATGAFANKDQOEKHAKEATEALENIRTVSLTR 830  
 896 FNMKSLIVISVFPFLALSGAVQTKMLTGSPQSKELLEKAGQILNEALSNIRTVAGICV 955  
 831 EKAFQOMTEEMIQTOHRNTSKKAQIIGSCYAFSAFIYFAVYAGRFPGAYVLIQGRMPE 890  
 956 EGRFTKAEVELEKSYKAIKRVANYGLCYAFSGISFLANSAAVRYGYLLIVYEDLANS 1015  
 891 GMFIYFPAIAGAMAIIGTVLAPYSKAKSAAHLFALLEKPRINDISQSGKRPDCE 950

DB 1016 YFRRVSSIAMSATAVGRFSTPYAKKISAARFOLDLKKPPIDVYSSAGEKMDNQ 1075  
 QY 951 GMLFEVRSFFYPGRDPFVILGSLISIERGKTVAFAVSSGGCKSTVOLRLYDPVOG 1010  
 DB 1076 GKIDEDICKFTFYPSPDIOVLNGLSVSDPGOTLAFVSSGGCKSTIOLLERFYDDPG 1135  
 QY 1011 QVLPFGVAKELNVQMRQIAIVPOEPLFNCSTIAENIYAGDNSRVVPLDIKEAANA 1070  
 DB 1136 TWMIDGHSKRVNVOPLKRNIGIVSOEPLFDCSTINDIKYDNTKEISVERAIAAKOA 1195  
 QY 1071 NISFTEGPEKYNFQVGLKGAQLSGQKORLAIRALLOKPKILLIDETSALDNDSEK 1130  
 DB 1196 QAHDEVMSPEKYEFTNNGVGSQSLRGEKORIAIRAIYRDRKILLIDATSALDTESEK 1255  
 QY 1131 VVQHALDKARTGTCVLYTHRLSAIONADLYLVHNGKIKEGTQHOELLRNDRDYFKLV 1189  
 DB 1256 TVQALMDKAREGRTGCIIVHRLSTIONSNDIIVMSQGVIEKGTGKHLMDQKAYKLV 1314

## RESULT 14

AB11\_RABIT STANDARD; PRT: 1321 AA.  
 AC Q9NOV3:  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)  
 DE (SISTER OF P-GLYCOPROTEIN).  
 GN ABCB11 OR BSEP OR SPGP.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;  
 RA Balasubramanian N.V., Suchy F.J., Ananthanarayanan M.;  
 RT "Molecular cloning and characterization of rabbit liver bile salt  
 RT export pump (Bsep/spgp).";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 CC -I- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
 CC INTO THE CANALICULUS OF HEPATOCYTES.  
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
 CC SITU.  
 CC -I- DOMAIN: MULTIFUNCTIONAL. POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
 CC ATP BINDING CASSETTE (ABC) DOMAIN.  
 CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF249879; AAF65552.1; -.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001140; ABC\_transport\_tmem.  
 DR InterPro: IPR001687; ATP\_gmp\_a.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran\_2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transmembrane; Transport.  
 FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).

FT	TRANSMEM	63	83	POTENTIAL.	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	84	147	POTENTIAL.	POTENTIAL.
FT	TRANSMEM	148	168	POTENTIAL.	POTENTIAL.
FT	DOMAIN	169	215	CYTOPLASMIC (POTENTIAL).	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	216	236	POTENTIAL.	POTENTIAL.
FT	DOMAIN	237	240	EXTRACELLULAR (POTENTIAL).	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	241	261	POTENTIAL.	POTENTIAL.
FT	DOMAIN	262	319	CYTOPLASMIC (POTENTIAL).	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	320	340	POTENTIAL.	POTENTIAL.
FT	DOMAIN	341	353	EXTRACELLULAR (POTENTIAL).	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	354	374	POTENTIAL.	POTENTIAL.
FT	DOMAIN	375	755	CYTOPLASMIC (POTENTIAL).	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	756	776	POTENTIAL.	POTENTIAL.
FT	DOMAIN	777	794	EXTRACELLULAR (POTENTIAL).	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	795	815	POTENTIAL.	POTENTIAL.
FT	DOMAIN	816	869	CYTOPLASMIC (POTENTIAL).	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	870	890	POTENTIAL.	POTENTIAL.
FT	TRANSMEM	891	911	POTENTIAL.	POTENTIAL.
FT	DOMAIN	912	979	CYTOPLASMIC (POTENTIAL).	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	980	1000	POTENTIAL.	POTENTIAL.
FT	DOMAIN	1001	1011	EXTRACELLULAR (POTENTIAL).	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1012	1032	POTENTIAL.	POTENTIAL.
FT	DOMAIN	1033	1321	CYTOPLASMIC (POTENTIAL).	CYTOPLASMIC (POTENTIAL).
FT	NP_BIND	462	462	ATP (POTENTIAL).	ATP (POTENTIAL).
FT	NP_BIND	1113	1120	ATP (POTENTIAL).	ATP (POTENTIAL).
FT	CARBOHYD	109	109	N-LINKED (GLCNAC. . .)	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	116	116	N-LINKED (GLCNAC. . .)	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	122	122	N-LINKED (GLCNAC. . .)	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	125	125	N-LINKED (GLCNAC. . .)	N-LINKED (GLCNAC. . .)
SO	SEQUENCE	1321	AA: 146376	MM: 457539FC06D17A2	CR664:

Query Match	46.0%;	Score 2765;	DB 1;	Length 1321;
Best Local Similarity	44.1%;	Pred. NO. 2.1e-153;		
Matches 555;	Conservative 243;	Mismatches 394;	Indels 66;	Gaps 6;

```

0Y 1 MIIGLASIVNGACPLMPPLVLGSEMDLI-----SGCLVONTY-----40
Db 62 MCMGSLCACIHGIAOPGVLLTGTMTDVFIDYDPELOELKTPGKACNNNTIWNSSLNQ 121
0Y 41 -----SPFRLTLVYVGIGVAALIFGYIOISIMITTAAROTKRIKOPFHS 85
Db 122 NVTNGTRGGLDISEMIRFAGGYAIGISIVLTTGYIOWICWGTIAAHQLOKMKRSFRK 181
0Y 86 VLAODIGFDSODIGELNTRMF-DIDKISDGIKIALLFONMSTFISGLAVLWKWL 144
Db 182 IMRMGIGVAVDNCNSVKLNTPTSPSVFNKINDSSADQALFIQGMKSPJFGLFVGSOMKWL 24
0Y 145 TLVLTSTPLMASAASRNVILSTKSELASVSKAGVAEVLSSRTVYAFAGQKEL 204
Db 242 TLVITVSAPLLGLGAILIGLSVSKFTDELKAVAKAGSADEVISSMRTVAAPFGCKREV 301
0Y 205 QRYTONLMDAKDFGIKRTIASKVSIGAVYFPFNCTYGLAFWYGTSLINCEPGYITVL 264
Db 302 ERYENNLVFAQRMWGRKRIYWGFTGYMWCILFFCYALAFYSGKLYE-BGEYSPPALV 366
0Y 265 AVFSEVHSYVCIGAAPHFTFALARGAAPHIFQVIDKPSIDNFSTAGYKPPESIGTV 324
Db 361 QIFLSVIGIALNGLNASPCLEAFAGRAAASSIFETIDRKPJIDOMSBDGKLERINGEI 426
0Y 325 EFKNFSFVYPSRPSKILIKGLNLRKSEYIALVGLNGSGKSTYVQLRLYDPDDOFIM 384
Db 421 EFHNTFPHYPSRPEVKILNNLSMYIKPEMVALPGSAGKSTALQILHREYGPTECMVT 480
0Y 385 VDENDIRALNRHYVDHIGVAVSOEVLFGTTISNNIKRGPDVADDEMERARARANYDF 444
Db 441 VESHDIRSHLOMLRNQIGIYQEPVLEFTHIAEKIRIGREDATMEDJIOAKRANNYNF 540
0Y 445 IMEPKNCNTLVGEKGAOMSGOKORIALAVALRNPKIILLDEATSAIDSESSESAVOAA 504
Db 541 IMDLPQGDVLVGEKGGOMSGOKORIVALAIALRINRPIILLDMATSAIDNESAMVOEA 600
0Y 505 LEKASKGRTTVVAHRUSTINSADILVTLKQDMLEKGAHAELMAKRGLYISLVMSODIK 564

```

Db	601	LSKQHCHTIVSAHNRATIRTDVYIGEHOAAVREGEELERKVYALVTLQSR	660
Qy	565	KADEQESMTYSTE-----RKTNLSJPLHSVKSISDPLDK	599
Db	661	NOGOEENEDATEDDIPEKTFESRGVYODSLASLRORSKOLSYLAHEPPMAVEDHKST	720
Qy	600	AEESTQKESLDE---VSLKILILNKPENPVLVIGLASLNGVHPAFESIIFAKII	655
Db	721	HEEDRKRDJLPAQODIEPPASVRIRIMLNPEWYIMMLGSMGAANVATPLLYALFESOL	780
Qy	656	TMEGNNDKJTLTKHDAEISYMEFVLIGVIFSVYFMQGLFEYGAGEILTMIRLHAFKAML	715
Db	781	GTESLPEKEQROQINICGLFVTLQCSFEPFOQYFPAASGILLRLKRKGFRML	840
Qy	716	YODIAMFDEKENSTGGITTLTILAIDIAIOGATGSRIGVLTONATMGLSVIISFTYENEM	775
Db	841	GODIGMFDDLRNSRGATLTTLATDASQVAGATGSOIGMVMNFEYVYAMIAIAFLFSMKL	900
Qy	776	TFELLSTAPLAVATGMETAAMGFANRKOELKHAQKATATEALEINIFTIYSLRKAQE	835
Db	901	TLGVICFEPPLASLGALQJTMLTGFSRKOALERKQJTSLSINITVYAGIGERKEFI	960
Qy	836	OMYEEMJOTQHRMRTSKKAQITISCSYAFSHAIFYFAVAGREFGALLIOGAMTEGMEIV	895
Db	961	ETFEAELEKRYKMAIKKANYGIGLCFEGSOCIIFIANSSARYCGVLLINEGJHSYSYRV	1020
Qy	896	FTAIYAGAMAIGTLVLAPYSAKSGAANHLPELLEKKNPINDRSQBECKKPDTCBNLEF	955
Db	1021	ISAVVLSATLGRASSYTPPYAKAKISARFOLLDRQPIWVYSSAEKMDNFCQKIDF	1080
Qy	956	REVSFFPCQPDVFLILGRSLSTERKTYAFNGSSCGCKSTSVOLQTLQYOPGYOVLFD	1015
Db	1081	VDCKEYTPSPRPDIQVINGLSVSPSPROTALFEGSSCGCKSTSIQLLEKRYPDDBEKWID	1140
Qy	1016	GVADKEINVOMLQSOLAIYQOEVLVFNCSIAENIAYGDNRSRVPLDEKREKAAANAHISF	1075
Db	1141	GHDKRKNIGQIFLKSNGICVYQSEPVFLFACSTIKDNIKYKGDQTCIIPMERIITAANKAQYHDF	1200
Qy	1076	IEGLPERKYNTOVGLKAGLSGGOKORALAIARALLQPKTLLIDEATYSAUNDSEKVOVQA	1135
Db	1201	VMSPEKYEYTNVSGQSLSRGEKORITAIRAIVRDPKILLDEATSAJLDESEKTYOVA	1260
Qy	1136	LDKARQRTGLVYTHRLSALQNMDDLVLHANKIKYBOGHQHELLANRPYKINAVNAGS	1193
Db	1261	LDKARBERTIVIAHNRSTJQNSDIIIAVWSQGVYIEKGEHLEMAQKAYTKVLVTGS	1318

MDRL_15			
MDRL_CAEEL			
ID	MDRL_CAEEL	STANDARD:	PRT; 1321 AA.
AC	P34712;		
DT	01-FEB-1994 (Rel. 28, Created)		
DT	01-FEB-1994 (Rel. 28, Last sequence update)		
DT	01-FEB-1994 (Rel. 28, Last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN A).		
GN	PGP-1.		
GN	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Palpodoineae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RX	MEDLINE=93085750; PubMed=1360540;		
RA	Lincke C.R., The I., van Groenigen M., Borst P.;		
RT	"The P-glycoprotein gene family of Caenorhabditis elegans. Cloning		
RT	and Characterization of genomic and complementary DNA sequences."		
RL	J. Mol. Biol. 228:701-711(1993).		
RN	[2]		
RP	TISSUE SPECIFICITY.		
RC	STRAIN=BRISTOL N2;		
RX	MEDLINE=93223702; PubMed=8096815;		



RA Lincke C.R., Broeks A., The I., Plasterk H.A., Borst P.;  
RT "The expression of two P-glycoprotein (pgp) genes in transgenic  
RL Caenorhabditis elegans is confined to intestinal cells.";  
EMBO J. 12:1615-1620(1993).  
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: INTESTINAL CELLS.  
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X65054; CAA46190.1; -  
DR PIR: S27337; S27337.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001140; ABC\_transportr\_tmem.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
KW Multigene family.  
FT DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 78 98 POTENTIAL.  
FT TRANSMEM 144 164 POTENTIAL.  
FT TRANSMEM 213 233 POTENTIAL.  
FT TRANSMEM 240 260 POTENTIAL.  
FT TRANSMEM 321 341 POTENTIAL.  
FT TRANSMEM 350 370 POTENTIAL.  
FT DOMAIN 371 753 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 754 774 POTENTIAL.  
FT TRANSMEM 798 818 POTENTIAL.  
FT TRANSMEM 874 894 POTENTIAL.  
FT TRANSMEM 895 915 POTENTIAL.  
FT TRANSMEM 978 998 POTENTIAL.  
FT TRANSMEM 1017 1037 POTENTIAL.  
FT DOMAIN 1038 1321 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 451 458 ATP (POTENTIAL).  
FT NP\_BIND 1112 1119 ATP (POTENTIAL).  
SQ SEQUENCE 1321 AA; 145074 MW; 6130A9F3B31A8FA9 CRC64;

Query Match 38.7%; Score 2326; DB 1; Length 1321;  
Best local similarity 40.4%; Pred. No. 8.4e-128;  
Matches 504; Conservative 225; Mismatches 462; Indels 56; Gaps 10;

QY 1 MTLGLASLVNACPIPLMPLVIGEMSDNISGLVQNTYSPF----- 43  
DB 77 LFTIGLVAVITAGLPLMSILOGKVSQAFINQIVAINNNGSFTLPTGQNTYRTDEHDVY 136  
QY 44 RLTYLVYVIGVALLFGYQIQISLITTAARQTRIKRQFHFVLAODIMFSDIGELN 103  
DB 137 NVVYSYAMTVMGMAAGQIVTCYLVAGQNNRLRREVKSILNGEISMPTNHSGLTA 196  
QY 104 TRMTD-IDKISGIDDKIALTLFQNNSTESIGLAVGVKWLTLVTLSTPLMASAAC 162  
DB 197 TLFPLNLEKVEKGTGDKIGMAFOYLSQFTFTGFIATFTHSMTLVLAVTPIQALCGFAI 256  
QY 163 SMVVISLRSKEISAYSKAGVAEVLSSRTYIAFPAQKEQORTQNKDAKDFSIKRT 222  
DB 257 AASMSGFAIRELTRYAKAKAVVEETISIRTVSLNGLRYELERSTVAEEKAKGAVLNG 316  
QY 223 IASKVSIAGVYFPMNGTYGLAFWYGTSLINCEPGYITGVLAFFSVYHSSYICGAAP 282

DB 317 LFLGISFGMAQASNFISFALAFYIGVWHDGSLNF--GDMLTFESSVMGSMALGLAP 374  
QY 283 HETFAIANGAFHIFQVIDKRPSIDNFTAGYKPESTIGYEFKNVSNVRSRSIKTL 342  
DB 375 QIAVIGTAGAAGSAGIYEVDKRPVIDSSSKARCKDKIKGDITVEVHTYPSRDVPL 434  
QY 343 KGLNTRIKGQEVVALVGLNGSGSKSTVVOQLRYLPPDDGFIMWENDIDIALVVRHDI 402  
DB 435 RGMNLRVANGQVALVGVSSGCKSTIISLLRYDYVLRKTIITDGDVDVINDLEFLRKVY 494  
QY 403 GYVSOEPLVLEGTTSINNIKKYGRDVTDEEMERAREANAYDEIMEPNKENTLVGEKGAQ 462  
DB 495 AVYSEPALFNCTIEENISLGEKGTIREMVAACKMAKEFKTLPLNGYNTLVGRGTQ 554  
QY 463 MGGGOKORAIARALVRNKKIILDEATSAIDSEKSAVQAALERASKRTITVAHRIS 522  
DB 555 LSGGOKORAIARALVRNKKIILDEATSAIDSEKSAVQAALDRNAKRTYIIAHRIS 614  
QY 523 TRSADLIYTLKDGMLAEKGAHAEIARGLYSLVMSOD-----IKKADEQMSMTYTE 578  
DB 615 TIRNADLIISCKANGQYVEGDHRAIQAQGLYDLYTAOTFTDAVDAEAGKFSRNSVA 674  
QY 579 RKT-----NSLPLHSYKSI-KSFIDRAE-----STQSKETS 610  
DB 675 RQTSHEGLSRQASEMDIMNVRSSSTIGTNGPYIDEKEERIGKDALSRUKOLEENN 734  
QY 611 LPEVSLIKTLKLNKPEWPPVYVGLTASVLANGTVHPVFSIIIPAKIITFEQNNKTKIKHA 670  
DB 735 AOKTNLFEILYHARPAHLSLFTIGMSYATIGFTYPPYSVFVFSFNVFAGNADFLS-OG 793  
QY 671 EYISMFEVLGYICFVSYMOGLFYGRAGEIITMRRLHIAFAFAMLYQDIAFMEKENSNG 730  
DB 794 HFWALMEFLVLAAGCISLMTFPMGIASESITRDILRNKLFRNVLSOHIGFDSQONASG 853  
QY 731 GLTTLAIDIAIOGATGSRIGVLTQMATNMGISVIISPIYGEWMTFLLISAPVLAVTG 790  
DB 854 KISTRLATGVPLRFAIDRFSTVITLVSVAGIGLAFYGVQMAILLIATLPIYAFQO 913  
QY 791 MIETAMTFANKDKOELKHAKGIATELENTIYSLTRKAFEDQYEMLOTOHRNRS 850  
DB 914 YLRGRFETGKNVKSASEFDSGKIAIEAENVTQVALREDETFEYENPECKIDIPKKA 973  
QY 851 KKAQIIIGSYARSHAFIYAVYAGPFRGAVLT--QAGRMTPGSMFTYFPAIYAGAAAIK 908  
DB 974 KEAFTQGLSYGASSVLYLNTCAVYMGALITTDPTMQPMRVLRVMTAITISTTLGF 1033  
QY 909 TVLAPEYSKAKSGAAHLFALLEKKPNIDRSQOEGKAPDTCEGNLEFRVSEFFYPCRPDV 968  
DB 1034 ATSYPEYKAKFPAGGIIIFGMRLKISKIDSLISLAGEK-KKLGVKVIKFNVRAYPERPEI 1092  
QY 969 FTLRGLSISIERKTYAVFYVSSGCGKSTSVOLLRLYDPVQGVLFQVDAKELNVQMLR 1028  
DB 1093 ELTKGSLFSEVGEQATLALVPGSCGKSTYVALLEERYDYDLGEIIFDGEISLTLNBEHR 1152  
QY 1029 SOLAIYPOEPLVFNCISIAINIAIYGDPSVYMAQVEEARLNIHFIAELPEGFETRQ 1088  
DB 1153 SOLAIYQEPTEFDSIANIITIGDPSVYMAQVEEARLNIHFIAELPEGFETRQ 1212  
QY 1089 LKGAOLSGGOKORAIARALQPKILLDEATSAIDNSENKQVYVQALDARFGRYCLV 1148  
DB 1213 DRGTOISGQOKRIARALVRNPKILLDEATSAIDSEKSAVQAALDARARGRCIYI 1272  
QY 1149 THRSLAIONADLIYLVHNGKIRKQGTQHELNRNDIYFVLVNAOSVO 1195  
DB 1273 AHRLNTVMNADCIAYVNSGTIIEKGTHTQDMSEKGAVYKLTOKOMTE 1319

Search completed: April 22, 2002, 14:37:38  
Job time: 290 sec

---





Db 50 MLIGTLAIIHGIALPLMLVFGMDTDSFAN--VGNRRSNSEFYNAVDYAKLEDENITVA 107  
 QY 47 LYYVIGVALALGYIOISLWITTAARQTRIKRQPFHSLVLAODIGWDCDIDGELNTRM 106  
 Db 108 YTTTGAGVLLIAYIYQVSWCLACRQIHKIKQFFHAIIMNOIGWFDVHDELNTRL 167  
 QY 107 T-DIDKISDGIKLLDQNMSTFSGIAGLVKGMKLTLYLSTSPILMASAACSRM 165  
 Db 168 TDDVSKINIEGDKIMFQAMATPFQGFITGFTRGKLTLYLALISPVLSAGIMAKI 227  
 QY 166 VISLTSKELSAISKAGAVAEVLSIRTYAFAPAOEKLORYONLKDADFGIKRTIAS 225  
 Db 228 LSEFTDEKLOAYAKAGAVAEVLAALRTYAFAGOKKELERYNNNEAEAKRLIKAITA 287  
 QY 226 KVELGAVYFPMNGTYGAFWYGTSLINGBPGYTIGTVLAVFSSVSSYICGAAYHFE 285  
 Db 288 NIMGAFLIYASYALAFWYGTSLVSKR--YTIQGVLFVFSYVLGASVQOASNIE 345  
 QY 286 TFAIANGAFHIQVIDKRPISDNFSTAGYKPSIEGTFEKNVSENPSPSIKILKGL 345  
 Db 346 AFANARGAAYEFSIIDNKPSIDSFSGSKHKNIOGNLEFKNIHFESYPSRKDVQLKGL 405  
 QY 346 NLRKSGEYALVGLNSGKSTYVOLLQRLYDDDFIMVDEMDIRLANRHRDHTGYV 405  
 Db 406 NLKVGSGOYVALVNSGSGKSTYVOLLQRLYDPEGEVSIDGDIRINRYLREIIGVY 465  
 QY 406 SOEPVLFETTISNNIKYGRDVTDEEMERARAPANATDFIMEFPNKENTYVGEKAQMSG 465  
 Db 466 SOEPVLFATTIAENIRKGRNVMTDELEKAVKENAVDFIMKLPKRKGDVLGEGQALSG 525  
 QY 466 GOKORIAIARALVNRPKILILDEATSALESSESAVOALEKASKGRTTYVAHRLSTIR 525  
 Db 526 GOKORIAIARALVNRPKILILDEATSALESSESAVOALDKAREGRTTYVAHRLSTVR 585  
 QY 526 SALLIYTLKGMALAEKGAHLEMAKRGLYYSYVMSQ----DIKKADQOMS-----MTY 575  
 Db 586 NADVIAGEFDGVLVEQGNHDELMREKGIYKLVMTQTAGNEIDENACSDGODINDVM 645  
 QY 576 STERKTSNL--PLHSVYSIKSDFIDKAEEST--OSKEISLEVEVSLKILKINKRPMPEPV 632  
 Db 646 SSKDSSGSLLRSTKRSKISPHDQOGELESTKEALDDVDVPASWRLILKLNSTMPPEFV 705  
 QY 633 GTIASVINGTVHPVFSIIFAKIITMFGNNDKTYL-KHDAEYSMIEVILGVICFVSYFMQ 691  
 Db 706 GVECAIINGLOPAFSIIFSGVGVFTKNDPEIORONSLFSLFLILGISITPEFIQ 765  
 QY 692 GLEPGRAGELTMTLRRLAKRMLYODIAMPDEKENSTGGTTLADIDIOIGATSR 751  
 Db 766 GFTFGKGEILTKRLRYMFKSMLRQDISWDDPKNTTGALTTRLANDAQAQKATGSR 825  
 QY 752 GVLTONATNNGLSVYISFIYGMETPELILSIAPVLAVTGMIETAAMTGFANKDKQELKHA 811  
 Db 826 AVITONANIGTGIILISLIIGWOTLLALLAIVPILAIAGVEMKMLSGQALDKKKELEGS 885  
 QY 812 GKATEALENIRITVSLTREKAEOMTEMLQTOHNTSKKAOITGSCYAFSHAFTYFAY 871  
 Db 886 GKATEALENIRITVSLTREKAEOMTEMLQTOHNTSKKAOITGSCYAFSHAFTYFAY 945  
 QY 872 AAGRFAVYLIQGRMPEGMFIYETAIAAGAMAGIKTIVLAPREYKASGAALFLALLE 931  
 Db 946 AACRFAYVILVARELMFEVNLVFSALVFGAMAVGOVSSFADPYAAKAKYASHIIRIE 1005  
 QY 932 KKPIDRSOEGKPKDCEGNLEPREVSPFPKPPDVEILFGLSLSTJERGTVLAVFSSG 991  
 Db 1006 KIPIDISYSEGLKPNMLEGNVKNFNGVMPNTPRPNIPVLOGLSLEVKKQOTLAVSSSG 1065  
 QY 992 CGKSTVOLLQRLYDVOGVLFEDGVDAKELANVOMLSQIATVPOEVLVFNCSIAENIAY 1051  
 Db 1066 CGKSTVOLLQRLYDVOGVLFEDGVDAKELANVOMLSQIATVPOEVLVFNCSIAENIAY 1125  
 QY 1052 GDNRRVYPLDEIKAAANANHSFTIEGLPEKYNTVOVGIKGAQLSGGOKORIAIARALLQK 1111  
 Db 1126 GDNRRVYSHIEIVAKAEANIHOFIDSLPEKYNTVGDGKQTLSSGGOKORIAIARALLQK 1185

QY 1112 PHILLIDEATSALENDSEKVVQHALDKARTGTCVYTHRLSAIQNADLLVLANKIKE 1171  
 Db 1186 PHILLIDEATSALENDSEKVVQHALDKARTGTCVYTHRLSAIQNADLLVLANKIKE 1245  
 QY 1172 QGTHOELNRNDIYFKLVNAOS 1193  
 Db 1246 HGTTHOOLAKGIGYFSNVQVA 1267  
 RESULT 2  
 002793  
 ID 002793 PRELIMINARY; PRT; 1285 AA.  
 AC 002793;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN-1.  
 GN MDRL.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Longley M., Crawford A.M.;  
 RC Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS)  
 CC EMBL: U78609; AAB58489.1; -.  
 DR Interpro: IPR003593; AAA.  
 DR Interpro: IPR001140; ABC\_transporter\_tmam.  
 DR Interpro: IPR003439; ABC\_transport.  
 DR Interpro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PR000664; ABC\_membrane\_2.  
 DR Pfam: PR000005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KM ATP-binding; Transport.  
 SQ SEQUENCE 1285 AA; 142020 MW; 90153E617C44856F CRC64;  
 Query Match 59.0%; Score 3549.5; DB 6; Length 1285;  
 Best Local Similarity 55.5%; Pred. No. 7.5e-197;  
 Matches 685; Conservative 237; Mismatches 262; Indels 51; Gaps 9;  
 QY 1 MIIGTILSVNGACLPLMPVLYGEMSDNLIS---GCLVQTN----- 38  
 Db 55 MVLGTLAIIHGIALPLMLVFGMDTDSFAGAGNLGNTLSNSTIDRTEYKKLEK 114  
 QY 39 ---YSEFRLTYVVGVALALFYIOISLWITTAARQTRIKRQPFHSLVLAODIGMFD 95  
 Db 115 EMITTYA-----YSGIGAVLLIAYIYQVSWCLACRQIHKIKQFFHAIIMQOIGMFD 169  
 QY 96 SCDIGELNTRMT-DIDKISDGIKLLDQNMSTFSGIAGLVKGMKLTLYLSTSP 154  
 Db 170 VHDVGKINTRLTNDVSKINIEGDKIMFQAMATPFQGFITGFTRGKLTLYLALISPV 229  
 QY 155 IMASNAACSRMVLISLTSKELSAISKAGAVAEVLSIRTYAFAPAOEKLORYONLKD 214  
 Db 230 LGSAAIATWAKLSFTFTRKLLAAKAGAVAEVLAALRTYAFAGOKKELERYNNNEAE 289  
 QY 215 KDFGIRRTIASKVLGAVYFPMNGTYGAFWYGTSLINGBPGYTIGTVLAVFSSVSHSS 274  
 Db 290 KRIGIKRALTAINISMGAFLLIYASYALAFWYGTSLVSKR--YTIQGVLFVFSYVLG 347  
 QY 275 YCIGAAVPHETFAIANGAFHIQVIDKRPISDNFSTAGYKPSIEGTFEKNVSENP 334  
 Db 348 FSIQASPNIEAFANANGAAAYEFSIIDNKPSIDSFSGSKHKNIOGNLEFKNIHFES 407  
 QY 335 SRPSIKILKGLNLRKSGEYALVGLNSGKSTYVOLLQRLYDDDFIMVDEMDIRLAN 394

Dh	408	SRNEVTKILKGLNKLKVGSGQIVLALVNGSGCKSTTVYLMRLYDPIEGWMSIDGODIRIN	467
Qy	395	VRHYRDIIGVAVSEPEFLGTTISNNIKYGRDVTDEEMERARAANAADYDFIMEFPKNFT	454
Dh	468	VRYLRREITIGVSGPEVLFAATTIAENIRYGRDEDVTMDIQKAVKEANAADYDFIMKLPNKFDI	527
Qy	455	LVGRKGQMGSGGQKORAIARALVRNPKIILDEATSALDSKSKAVQAALAEKASKRTT	514
Dh	528	LVGRGQALSGGQKORAIARALVRNPKIILDEATSALGTESEAVVQAALDAKARKGRTT	587
Qy	515	IVVAHRSTRSADLIYTLTADGMAELAEKAEALMAKRGILYSYVMSQDKKXDEQMEST	574
Dh	588	IVIAHRSTARNADITAGLDDGVIVERGSHDELMKRGITFLVLYMQ-TKGELELENTP	646
Qy	575	YSTERKTNLSPLHVSXIKSDFI-----DKAESSTQSKELSLPEVSLIKI	619
Dh	647	GESLSTKIDDLVMSQDS-RSSLIRKSTRSIRNGSQDQRKLTSTEBTLDSEVPVPSFRI	705
Qy	620	LKLKPEMPVVGTLASVINGTVHVPFSTIEAKITIMFGNNDK-TTLKHDAEISMTIV	678
Dh	706	LKLTITEMPVFVVGVCFAIINGALQAPAFSYFIRIGITFRINDNETKRONSLFLFL	765
Qy	679	ILVGYCVSYFMGCLFTGRGCELLTMRRLNARKALAYODIAMFDEKESNGTLLIAT	738
Dh	766	ILGIISITFEFLGQFFPKAGELLTRLRLVAFRSLRDVSWFDDPKNTTALTALTRLAN	825
Qy	739	DIAIOGATSRIGVLQONTNGLSITISFIYGMFPLIISAPVLAVTGMETAMT	798
Dh	826	DAAVKAQAVSRLLVITQNTANLGTGIIISLIGMOLTLTLALVPTIVAGVIEKMLS	885
Qy	799	GFANKDKOELHACKIATEALENIIRTIVSLTREKAEQWYEMELOTHRNTSKAKIIGS	858
Dh	886	GOALKDKKELEGACKIATEALENFRIVSLTREBEREYWAQSLQVPRANSLRAHVPFI	945
Qy	859	CYAFSAHFYFPAVYAGRFCAATYIQGRMTPEGMTVFYPAIAGAMAIKTYLAPETSK	918
Dh	946	TFSTIQTQAMFYFVSGCFRFAATYVAGIIMFQDVLTVFSAVVFGMAAVGSTFAPDYAK	1005
Qy	919	AKSGAHLFALLEKPKMIDRSOEGKKPDCBCENLFFRVSFEPYPCRPVPLTRGISTSI	978
Dh	1006	AKVSAHAHINITEKIPILDSITSTEGKLPSSTVBESVAFNDVWFYPRPVPVLRGLSTLV	1065
Qy	979	ERGRTAFFVGSSGGCKSTSVOLLQRLYDIPVQGVLDGDAKELNAYQMLRSOIAIVPEP	1038
Dh	1066	KKGGTLALVSGSGCKSTVQVQLLERFYDPLAGVFLIDGREVQNLANYQMLRAMGVISGP	1125
Qy	1039	VLFNCSTIAENTAVGDSNRVYPLDEIKEAANAANISHFISGLPEKTYTQVGLKGAOLSGQ	1099
Dh	1126	ILFPCISIGENTAVGDSNRVYVSOEIEHAAKEANISHFIMLPDKYTRVGDGTGTLSGQ	1185
Qy	1099	KORLAIRALLOKRIILLDEATSALDONSEKVVQALDKARTRGRCLVYTRLSAQNA	1155
Dh	1186	KORLAIRALROPHIILLDEATSALDTESEKVVQALDKARGRCTIVIAHRISTONA	1245
Qy	1159	DLIVVLHNGKIKEGQTHQOELLNRNDYFXYKLVNQS	1193
Dh	1246	DLIVVPONGRIKEGHTHQOLLAQKGYITMYVQA	1280
RESULT 3			
046605		PRELIMINARY: PRT: 1280 AA.	
046605			
AC	046605:		
DT	01-JUN-1998 (TREMblrel. 06, Created)		
DT	01-JUN-1998 (TREMblrel. 07, last sequence update)		
DT	01-JUN-2001 (TREMblrel. 16, last annotation update)		
DE	MULTIDRUG RESISTANCE P-GLYCOPROTEIN.		
GN	MDR1.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TaxID=9615;		
RM	[1]		

```

RP      SEQUENCE FROM N.A.
RC      TISSUE=COLON.
RA      Puel O., Lepage J.F., Alvinerie M., Galtier P., Pileau T.;
RL      Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC      -i- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC      TRANSPORTERS).
DR      EMBL, AF045016; AAC02113.1; -.
DR      InterPro: IPR003593; AAA.
DR      InterPro: IPR001140; ABC_transporter_tmem.
DR      InterPro: IPR003439; ABC_transporter.
DR      InterPro: IPR001687; ABC_GMP_A.
DR      Pfam: PF00664; ABC_membrane; 2.
DR      Pfam: PF00005; ABC_tran; 2.
DR      SMART: SM00382; AAA; 2.
DR      PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW      ATP-binding; Transport.
SQ      SEQUENCE 1280 AA; 141523 MW; 762DD5AFAFC73306 CRC64;

Query Match          59.0%; Score 3549; DB 6; Length 1280;
Best Local Similarity 55.2%; Pred. No. 7,9e-197;
Matches 680; Conservative 239; Mismatches 270; Indels 42; Gaps

QY      1 MILGLASLVNCACLPLMLPLVGEKSDNLISGCL-----VQNTYSFFR--- 44
DB      50 MLVGTMAAIHGAALPLMLLVGMTDTSFANAGISRNKTFPVIIINESITFNTOHFINHL 109
QY      45 -----LTLYVGIVGAALIFGYIQSLMTITRAAOTKRIKQFHSVLADDIGMFDSCDI 99
DB      110 EEMTYYAYYSIGAGVLYAAVYIOVSFVCAAGAQQLIKRQFHAAIMRDEIGMFDVHDV 169
QY      100 GELNTRMT-DIDKISDGIGDKIALFQWMSFSGISGLAVGLKCGKMLLVLTSLSPILMAS 158
DB      170 GELNTRLTDVSKINIEGIGDKVGMFQSIATFFGTFVIGTFPGKMLLVLTSLSPVGLS 229
QY      159 AAACSRMYISLTSKELSAVSKAGAAVEEVLSSIRTVIAFRAQEKELQRYQNLKADAFG 218
DB      230 AAIMAKLISLSPDKELLAYAKAGAAVEEVLAAIRTVIAFGQKKELEBYRNMLEAKRIG 289
QY      219 IKRTASKVSLGAVFYFNANGTYGLAFWYGTSLINGEPGYITGVLAFFYSVHSSCTG 278
DB      290 IKKATTANISIGAAELLVYASALAFWGTSLVLSSE--YTIQGVLYFFVSVLGAQSG 347
QY      279 AAVPHFEFPAIRGAAPHFOYIDKPSIDNFSAGYKPEISIEGTVEFKNVFNYPSPS 338
DB      348 QASPSIEAFANRGAAYELFKIIDNKPSPIDSYSKSGHKPNIDIKNLEFKNVHSPSRKE 407
QY      339 IKILGLMLRIKSGETVALVGLNGSGKSTVYQLLQRLYDPDGFIMVDENDIRALNRY 398
DB      408 VKILGLNKLKVGSGTVALVNGSGCKSTVQLQRLYDPDGWVCIDGQIRITINRYHL 467
QY      399 RDHIGVSOEPLVFGITTSNNIKYGRDGYDDEMEARAENAYDFIMEPNKKNITLYGE 456
DB      468 REITGVSOEPLFAFTTIAENIRYGRENVYTDDELKAYKENAADFTMKLPNKEDTLVGE 527
QY      459 KQAKMSGGQKQIATARALVNRPKLLIDEPATSLDSEKSAVOALAKKSGKETTIVA 518
DB      528 KQARISGGQKQIATARALVNRPKLLIDEPATSLDSEAVVAYVALDKARKGTTIVIA 587
QY      519 HBLSTIRASDILLVTLLDKMLMEKGAHAEMLAKRGLYLSVMSODIKKADDEMESMTYTE 578
DB      588 HBLSTVIRANDVIAGRDGCVIYKGNHDELMKEKGIYFLVLMQ--TRNGEIELENAQESK 646
QY      579 KRTNSLPLHSVKSISDPT-----DKAESFQSKSEISLPEVSLIKITLN 623
DB      647 SESDIALEM-SPKDSOSSLSIKRSTRSRSTHAQOGDRKGTGEDINENVPSPVSFMRILKN 705
QY      624 KEMPFVYVGLTASVLNGVHVPESIIIFAKIITFM--GNNDKTTLKHAEIYSMTFVLGV 682
DB      706 STEMPFVYVGCALINGLOPANSIISRIGITFRDEDEETRKQNSNMFSVLFLVGI 755
QY      683 ICFVSYFMQGLFVYGRAGEITLMRLRHLAKFMKMLYODIAMPEKRNSTGGTLTIIADIAQ 742

```

Db 766 ISFTFFLOGFTGKAGELLTKRLRVNFRSMKODVSWEDPKNTGALTTRLANDAAQ 825  
Qy 743 IOGATGSRIGVLTONATNMGLSVIISFYGMEMTFLILSIAPLAVLAVMTAAMTFAN 802  
Db 826 VKGAGISRLAVITQNIANLIGIILISLIGMQLLLLLAIVPIAIAAGVEMKMLSQAL 885  
Qy 803 KDKOELKHAGKIATEALENIRTVSLTRKAEFEQYEMLOTOHRNTSKAQIIGSCYAF 862  
Db 886 KDKKELEGAGKIATEALENIRTVSLTRKAEFEQYEMLOTOHRNTSKAQIIGSCYAF 945  
Qy 863 SHAFIYAVYAAAFREGAVLLOAGRMTPEGMIVTATAIGMAIGKLVLAPEKSKAGS 922  
Db 946 TOAMAVYVAGCFREGAVLVAEPMFOVLLVSAIVFGMAAGVSSFAPDYAKAVS 1005  
Qy 923 AAHFALEKPPNDSDRSOEGKPPDCEGNLEFRESFFPYCRPDVLLRGLSLIRGK 982  
Db 1006 AAHYIMIEKSPLIDSYSPHCKRNTLEGANTFNEVFNPTRPDIYVQLSLSEVKKQ 1065  
Qy 983 TVAFVSSGCGKSTSVOLLQRLYDPVOGVLEVDVAKELNVMWLRSQLAIVPOEPVLEN 1042  
Db 1066 TLAVGSSGCGKSTSVOLLQRLYDPVOGVLEVDVAKELNVMWLRSQLAIVPOEPVLEN 1125  
Qy 1043 CSTAENIAYGDNRSVYVDELTEKEANANHSFIEGLPEKYNVOYGLKAGNLSGQORL 1102  
Db 1126 CSTAENIAYGDNRSVYVDELTEKEANANHSFIEGLPEKYNVOYGLKAGNLSGQORL 1185  
Qy 1103 AIAIALOKPKILLDEATSAALDNDSEKVVQAHADKARTGTCVLTHTLSAIONADLIY 1162  
Db 1186 AIAIALVROPHILLDEATSAALDNDSEKVVQAHADKARTGTCVLTHTLSAIONADLIY 1245  
Qy 1163 VLAHGKIRKOGTHOELLRNDRITFKLVNAQS 1193  
Db 1246 VFONGKVKHEGTHOELLRNDRITFKLVNAQS 1276  
RESULT 4  
ID 093437 PRELIMINARY; PRT: 1287 AA.  
AC 093437;  
DT 01-NOV-1998 (TREMblrel, 08, Created)  
DT 01-NOV-1998 (TREMblrel, 08, Last sequence update)  
DT 01-JUN-2001 (TREMblrel, 17, Last annotation update)  
DE ABC TRANSPORTER PROTEIN.  
GN CMDRI.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_taxid=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INTESTINE;  
RA Egelmann H.M.L., Duchek P., Rosenthal F.E., Poegeer N., Glackin C.,  
RA Kane S.E., Kuchler K.,  
RT Cmdri, a chicken P-glycoprotein, confers multidrug resistance and  
RT mediates estrogen transport.  
RU Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
-!- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
CC TRANSPORTERS).  
DR EMBL: AJ009799; CAA08835.1; -  
DR HSP; P13569; IMBD.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR InterPro: IPR000504; RRM.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR SMART: SM00382; AAA; 2.  
DR SMART: SM00211; ABC\_TRANSPORTER; 2.  
DR PROSITE: PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
KW ATP-binding; Transport.  
SQ SEQUENCE 1287 AA; 141785 MW; 3CFDBA8DD657111F CRC64;

Query Match 58.5%; Score 3519.5; DB 13; Length 1287;  
Best Local Similarity 55.0%; Pred. No. 4, le-195;  
Matches 673; Conservative 227; Mismatches 291; Indels 33; Gaps 5;  
Qy 1 MIIGTASLVNAGACPLMPVLGEMSDNL-----SGCVLTNTYS-----FFR 44  
Db 61 MIFSLAIAGHSLSPLTAMIIFGDMTDSFYSGMTNTTGNSSGINSADVFNKLEEMTR 120  
Qy 45 LTLVYGVAAALFEGYIOISLWITTAAROTKRIKQFPFHSVAQDQDSCDGLNT 104  
Db 121 YAYVYIAAALVAAVIAAIVTQSTFWTLAAGROYKTKREKFAIMQELGMEFDVNDAGLNT 180  
Qy 105 RM-IDIRKISDIDGDKTALLFPQNSTFSIGLAVLGVMKLTLLTSTSLIMASAAACS 163  
Db 181 RLIDVSKINGIDCKIFLQSTFTFLTGIVGFIFGKMLTVLILAVSVLGISALMA 240  
Qy 164 RMVISTSKELISAYSKAGAAVEVLSIRIVYIAFRAOEKELQRTOMLKADQFGIKRTI 223  
Db 241 KILFAFDKEQAAYAKAGAAVEVLSAVRVYIAFGCEKEIKRHKMLEPAKRIKIKAI 300  
Qy 224 ASKSLGAVYFENMGTYGLAFWGTSLINGEPYTGTVLAVFEVSIHSSYCIAGAVPH 283  
Db 301 TSNISMGAAFLIYASVALAWGTTLILANE--YISGNVLYFVSVLIGAFSIGTAPPS 358  
Qy 284 FEFPAIRGAFFHFOYIDKKPSIDNESTAGYKDESIETGVEEFKNVSNYPSRPSIKLK 343  
Db 359 IEAFANRGAAYAFNIIDNEPEIDSYSDAGHKRDHIKGNLEFQNVFNPSRPDEILK 418  
Qy 344 GLNLRISGETVALVGLNGSKSTVOLLQRLYDPDGFIMVENDIRALNVRIRRHIG 403  
Db 419 GLNLRVNGCGVVALVGLNGSKSTVOLLQRFYKPGCTTIDQDLKSLNVRILREIG 478  
Qy 404 VVSOEPLVFGTITSNNTKYGDVDTDEEMERAREANAAYPIMEFPKMFVLVEKQAM 463  
Db 479 VVNOEPVLFATTIENIRYGRDYMEIERATEANAYPIMLPKKEFVVEBERAQM 538  
Qy 464 SGGOKORIAIALVRNPKLILDEATSAALDSEKSAVAQALERASKGRITTVVAAHRLST 523  
Db 539 SGGOKORIAIALVRNPKLILDEATSAALDSEKSAVAQALERASKGRITTVVAAHRLST 598  
Qy 524 IRSADLVTLKOGMLAEKGAHAEIMAKRGLYSLVMSODIKKADEQMSMTYSTERKTN 583  
Db 599 VRNADLLAVFESGVITTEGNSHQLIEKKGIYKLVNMQTTETEDPSSEKSENAVSVRSG 658  
Qy 584 LPLHSVSIKSDF-----IDKAEESTQSEKISLPEVSLIKLKLPMPMP 629  
Db 659 SOSMLDESLEKLELRGSTRSMKRPGEPNQDTEKSGSPDELPVSLKLLKKNEMPY 718  
Qy 630 VILGTLASVLNGTVHPVPSIIFAKIIIMFGNNDKTTIKHDAEISMIVLIGVICFVSYP 689  
Db 719 FVATGFCALVAGLOPAFVSIFSEIIGFSETDKVLRKSNLSLFLAIGIISFTFF 778  
Qy 690 MGLFYGRAGITLRLHLAFKAMLYODIAMDPEKENSIGLTTIAlDAIOIGATGS 749  
Db 779 VQGFAGFGAGGILMKLRFMAFKMLRQDMAMPDPKNSGALTTRLANASOVKAGTGV 838  
Qy 750 RIGVLTQNTMGLSVIISFYGMEMTFLILSIAPLAVLAVMTAAMTFANRDKOKELK 809  
Db 839 RLALIAQNIANLIGIILISLIGMQLLLLLAIVPIAIAAGVEMKMLAGAKKDKIELE 898  
Qy 810 HAGKIATEALENIRTVSLTRKAEFEQYEMLOTOHRNTSKAQIIGSCYAFSHAFYF 869  
Db 899 AAGKIATEALENIRTVSLTRKAEFEQYEMLOTOHRNTSKAQIIGSCYAFSHAFYF 958  
Qy 870 AYAAGFREGAVLLOAGRMTPEGMIVTATAIGMAIGKLVLAPEKSKAGAAHFPAL 929  
Db 959 TYACGFRGAVLVNGHIEYTVFLVSAVYVFGAMALQGTSSFPADYAKAKISAAHFLVL 1018  
Qy 930 LEKPNIDRSOEGKPPDCEGNLEFRESFFPYCRPDVLLRGLSLIRGKLVLAPEKSKAGS 989  
Db 1019 FNRVPPIDSYREDEKPEKFGNTRIDVKFNPNRPEVKTLQGLNLAVEKGETLAVGS 1078

OY	990	SGCCGSEVSVOLQRJLYDPOGVGLDEVDGADELNVQWLRQSALAYPOEPFNCJIAENI	1049
		:    :                      :	
Db	1079	SGCGKSTVOYLLERYRYPDLSEGIYFDVDDIDAKTLNQMLRSHGIVSQEPILPDFITAEHI	1138
OY	1050	AYGDSNRVPPLDEIKEAAMAAINHSFIEGLEPEKYNTVOYLKGAGQLSGGOKORLAIRALL	1109
		:    :    :    :    :    :    :    :    :	
Db	1139	AYGDSRFVSHEEITSAAKAASIHFSIDLPEKYNTVRGDKGTQLSSGGOKORLAIRALI	1198
OY	1110	QKPRLILDEFTSALINDSEKVVOHALDKARTGRCLVYTHRLSAIQNADLIYVLHNGKI	1169
		:    :    :    :    :    :    :    :    :    :	
Db	1199	RKPOLILLDEFTSALDTSEKIVECALDKARGRCIYAHRSLTYIONAKDAVTIONGV	1258
OY	1170	KEOGTHOELLNRNDIFKLVNAOS 1193	
		:     :     :     :     :     :     :     :	
Db	1259	IEGTHOOLLAEKFYSISLVNVS 1282	
RESULT	5		
ID	060502	PRELIMINARY;	PRT: 1169 AA.
AC	060502;		
DT	01-NOV-1996 (TReMBUrel_01, Created)		
DT	01-NOV-1996 (TReMBUrel_01, Last sequence update)		
DT	01-JUN-2001 (TReMBUrel_17, Last annotation update)		
DE	P-GLYCOPROTEIN.		
GN	PGP-1.		
OS	Cricetulus griseus (Chinese hamster).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;		
OC	Cricetulus.		
OX	NCBI_Taxid=10029;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LUNG.		
RX	MEDLINE=91134265; PubMed=1671863;		
RA	Devine S.E., Hussain A., Davide J.P., Melera P.W.;		
RT	"Full length and alternatively spliced pgpl transcripts in multidrug-		
RT	resistant Chinese hamster lung cells.";		
RL	J. Biol. Chem. 266:4545-4555(1991).		
CC	-I- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC		
CC	TRANSPORTERS).		
DR	EMBL; M59254; AAA37005.1; -.		
DR	InterPro: IPR003593; AAA.		
DR	InterPro: IPR001140; ABC transporter_tmem.		
DR	InterPro: IPR003439; ABC transportr.		
DR	InterPro: IPR001687; ATP_GTP_A.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
KW	ATP-binding; Transport.		
SW	SEQUENCE 1169 AA; 128938 MW; 72E25B7CE29DC185 CRC64;		
Query Match	58.3%; Score 3506; DB 11; Length 1169;		
Best Local Similarity	56.6%; Pred. No.2.le-194;		
Matches	662; Conservative 238; Mismatches 248; Indels 22; Gaps		
OY	39	TYSFRRLLTVYGVAALIFEYIOISLMTITANROTNRKRQFFHSHVLAODIGMFDSCD	98
		:    :    :    :    :	
Db	3	TYAY-----YYTGAGAVLIAVYIOVSWCMLAAGRQIHKIRKPFPAIMNDEIGMFVDH	57
OY	99	IGELNTRMT-DIDKIDSIDGDKIALFPNNMFSFGVLGVGVKKLTLYLTSLPIMA	157
		:    :    :    :    :    :    :    :    :    :	
Db	58	VGELTRILTDVSKINEGIDGKITGMFQAARFEFGGITFGTRGKKTLTVIIAISVGL	117
OY	158	SAACSRNVISLTSKELSAVSRAGAVALVEVISSTFTVAIFAQOEKLORYTONLDKADF	217
		:    :                :    :    :    :    :	
Db	118	SAGIMAKITLSPTFKDLQAVAKAGVAEBVLAIRTVIAFGGQKKLELEYNNLEAPKL	177
OY	218	GKRIITAKVSLGANVFPMNGTYGLAFWTGYSLIANGEPGTYIGVLAVFESVISHSYCI	277
		:    :    :    :    :    :    :    :    :	

Dh	178	GINKATANISMGAFFLLIYASVALAFWJGTSJVSJSKE--YSIGQVLVVFVAUFIAPPSI	235
Qy	278	GAAPVHEFFETALARGAAFHIFOVYIDKPPSTIDNEFSTAGKRPESIBETVEFKVNSFNPSPR	337
Dh	236	GOASPNIIEAARGAALEYEFINIDDKPSTIDSESKXGKVPDNIKNGLNEFFKNHFSYPSRK	295
Qy	338	SIKIKGLNLRISGEFVALVYGLNSGKSGNVOLLORLVDPPDGGIYMDENDIRALNVRH	397
Dh	296	DYOLIKLNKLNVSGQIVALVNGSGGKSTYVOLLORLVDPPGVSISDGOIRIINRY	355
Qy	398	YRDIIGVSGOEPLFEGTISINNIKRYGRDVTYDEEMFARAEANAADFIMEFPNKENTLVG	457
Dh	356	LREIIGVSGOEPLFATTIENIRYRENTVYDEIEKAKENAAADFIMKLPHKRDVLVG	415
Qy	458	EKGAMSGGOKORAIARALVNRPKILLIDEATISALDSEKSAVOAALEKASKGRTTIV	517
Dh	416	ERGAQISGGOKORAIARALVNRPKILLIDEATISALDTESEAVOAAADKAREGRTTIV	475
Qy	518	AHRJSTIRSDLYTLTKDGLAEKGAHAELMKRGILYGLVNSQ-----DIKAD	567
Dh	476	AHRJSTIRVNDIILAGFGGIVYEGQHHEELMRKEGIYELVYTOFAGNEIEIGNEVSGSK	535
Qy	568	EOMSMYSPERKNTSL--PLHSYKSTKSDIFDKAESF--OSKEISLSEVSLTKTLKANK	624
Dh	536	NEIDNLMSSKSDASSLIRRRSTRKSIROGPHODKRLSTKEALDEDPYIPISWRILKANS	595
Qy	625	PEWPFVVLGTIASLVNGTVHVPFVSIIFAKITTF--GNNDKTLKHADEIYSMIFVLGY	683
Dh	596	SEMPYFVVGJFCAIVNGALQPAFISIFSKVGVGFRTNDETRKSDSNLSEILFILGVI	655
Qy	684	CFVSYFMOGLFYGRAGETILMRRLHAFKAMLYODIAMDEKENSNGITTLAIDINQI	743
Dh	656	SFIIFELQGTFFGAGELIKRLRYVFFKSMRKQDVSWPDNKNRTGALTITRLANDAGOV	715
Qy	744	QGATGSRIGVLTOMATMGSLVSIISFYGEMTEFLISIAPLVATGMETFAAMGFANK	803
Dh	716	KGATGALAVITTONIANLGIIISLILGHQTLILLALVPIIATAGVEMKMSGLK	775
Qy	804	DKOELKHAGIATEALENIRTIVSLTRERAKAFEDMYEEMLOTGHRNTSKKAOLIGSCYAS	863
Dh	776	DKRELBSGKIATEALENRTYVSLNREQEFEMYNQSLQIYRNALKKAAHYGITSFT	835
Qy	864	HAFIYFPAAGFPAGVLIIDAGMRTPEGMFYFTAIATYGAALIGTVLAPRYSKASGA	923
Dh	836	QAMMYFYSACFRGALVARELMEFENLVLFSAIVFGAMAVGVGSSFPADPYAKAKYSA	895
Qy	924	AHLELLEKKPNPIDSRSQBEKKKPDQECNLEPEVESEFPCPDVYILKGLSTISERKT	983
Dh	896	SHIIMILEKVPISIDYSTGKLKPNTELEGKNEVEYENPTTPBDIPLVLOGLEVEKKQOT	955
Qy	984	VAFVSGSGCKSTSVOLQLRLYDPVOGVLFDDVDAAKELVNOMLSQAIYVQEPVFLNC	1043
Dh	956	LALVSGSGCKSTYVOLLERFYPMAGTYFPLDCKEYNQOLNOMLRAHLGIVSQEPIFLDC	1015
Qy	1044	SIAGENIAYGNSRVPLDEIKEAANAANISHFETGLPERKYNTPOVLKGAOLSGGOKORIA	1103
Dh	1016	SIAGENIAYGNSRVQSODEIERAKKANIHOFIESLPDYKXNTRYVGKGTQOLSGGOKORIA	1075
Qy	1104	IARLLOKPKILLIDEATISALDSEKVVYOHADKARKRGTCGLVYTHRLSAIQNADLLVY	1163
Dh	1076	IARLVAQPHILLIDEATISALDSEKVVYOEALDKARBERTCIVIAHRLSTIQNADLLVY	1139
Qy	1164	LHNKIRKEOGTHOELLENRPYFKVLVNAS	1193
Dh	1136	IQNGKVEHGTHQOLLAGOKIYFSMYSVQA	1165
RESULT	6		
ID	Q91586		
AC	Q91586	PRELIMINARY:	PRT: 1287 AA.
DT	01-NOV-1996 (TREMblrel. 01, Created)		
DT	01-NOV-1996 (TREMblrel. 01, Last sequence update)		

DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN.  
 GN XEMDR.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;  
 OC Xenopodidae; Xenopus.  
 NCBI\_TaxID=8335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95322451; PubMed=7599185;  
 RA Castilho G., Shen H.J., Horwitz S.B.;  
 RT "A homologue of the mammalian multidrug resistance gene (mdr) is  
 functionally expressed in the intestine of *Xenopus laevis*.";  
 RL Biochem. Biophys. Acta 1262:113-123(1995).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS)  
 CC EMBL: U17608; AAA75000.1; -.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 KM ATP-binding; Transport.  
 SQ SEQUENCE 1287 AA; 141504 MW; 06695073C5771415 CRC64;

Query Match 57.6%; Score 3464; DB 13; Length 1287;  
 Best local similarity 55.0%; Pred. No. 6,7e-192;  
 Matches 673; Conservative 227; Mismatches 287; Indels 36; Gaps 9;

QY 1 MIIIGILASVNGALPLPLPLVIGMSDNLISGCLVQNTVSFFRL-----T 46  
 DB 66 MLFETIASLAGALPLMLVFGEMTDSFVNVGVDTGNFTWESMINASRELQOMTTYA 125  
 QY 47 LYYGIGAAALIFGIQISLWIITAAQRTKRIKQFHSVLAODIGMFDSCDIGNLNTM 106  
 DB 126 YYSGGLGFGVWLCAVIOISFTWLSAGROIKIRSNFHAIVRQELGMDINDADELNRL 185  
 QY 107 T-DIDKISDGIKIALLFQNMSTFSGILAVGLVGMKLTVLSTSPILMASAACSRL 165  
 DB 186 TDVSKINEGIGDKIAMLQSLTLVTGFIIGTIGMKLTVWMAISPIMLSAIAWAKV 245  
 QY 166 VISLSEKLSYKRAVAEVLSSIRTVAFRAQEKELQYTNQNLDAKFGIKRTIAS 225  
 DB 246 LSAFTNELKAYAKAGAAVEVLSSIRTVAFRAQEKELQYTNQNLDAKFGIKRTIAS 305  
 QY 226 KVSIGAVFFPANGTYGLAFWGTSLIINGEPYITIGTVLAVFFSVIHSYCIGAAPHFE 285  
 DB 306 NVSIGFAFLMIYAYSLAFWGTSLIINGEPYITIGTVLAVFFSVIHSYCIGAAPHFE 363  
 QY 286 TFAIARCAAFHIFOVIDKRPISIDNFSTAGVPESTIEGTFVEEKSNFNPSPRSIKIILKL 345  
 DB 364 AFANARGAAYTIFNIDQPKIDSEKSGKLKPKDKIDIEKFNVLFTYPSRKDIQVLKGL 423  
 QY 346 NLRIKSGEVALVGLNSGKSTVVOQLQRLDPPDGFIMNDENIDRALNVAHYHDHGV 405  
 DB 424 NLNIPSGTVALVGLSSGCKSTVVOQLQRFYDPDGVTLLDGRINIRINIRLEIIGV 483  
 QY 406 SQEPVLFCTSSNNIKYGRDQVTEDEMERARAEANAYDFIMEFPKFTLVGKGAQMSG 465  
 DB 484 SQEPVLFCTSSNNIKYGRDQVTEDEMERARAEANAYDFIMEFPKFTLVGKGAQMSG 543  
 QY 466 GOKORIALARLVNPKLILDEATSLDSEKSAVQALESKSGRTTIVAHRLSTIR 525  
 DB 544 GOKORIALARLVNPKLILDEATSLDSEKSAVQALESKSGRTTIVAHRLSTIR 603  
 QY 526 SADLVITLKDGMLEKGAHMAELAKRGILYSLWSODI---KKADEQMSMTYSERKTN 582  
 DB 604 NANAIAFGDNGVIVEOGSHKELMERGVYFNLVTLQTVETSKDELETHLY--EKRIIP 661

QY 583 SLPLHS-----VKSIKSDFDIK-AEESTQSKELSLPEVSLILKLKPKPEPV 630  
 DB 662 VTHTHSLVRRKSSRMTIKSKVPEDEDEEKKKEEPPVSEFKYKMLKPKPEPV 721  
 QY 631 VLGTSLAVLNGTYHPVPSITFAKITIMEGNNDKTLTKHDAEITSMIFVLIGVCFPSYFM 690  
 DB 722 VGVICAMINGATOPAPALLIFSRIGVFA-GPVSOHKSSEMTSLFLALGVGSFTTFEL 780  
 QY 691 QGLFYGAGELTMRRLHFLAKMALYODIAMPEKENSTGTLTTLAIDIAIOGATGSR 750  
 DB 781 QGLFYGAGELTMRRLHFLAKMALYODIAMPEKENSTGTLTTLAIDIAIOGATGSR 840  
 QY 751 IGVLTQATNMGSLVITISFYGEMFTLISLAVLAVTGMIFTAANTGANKDKDELH 810  
 DB 841 LALLAQVANMIGTAIIISFYGMQLTLLIATVAVIAAGLVEMKMPAGAKDKDELK 900  
 QY 811 AGKATATALENIRIVSLTEKAEQVEMLOHNTSKKAOIGSCAFSAFIYFA 870  
 DB 901 AGKISTDAVINIRIVSLTEKAEQVEMLOHNTSKKAOIGSCAFSAFIYFA 960  
 QY 871 YAAGFR-FGAVLIQAGWTEGMEFTVTAIAYGAMAIGTLVLAPEXSKAKSGAHLFAL 929  
 DB 961 LWFVSVLGAIVYEGMLKIDEVFLVSAIYLGAMALGQTSFPADPTKAMISAHIFSL 1020  
 QY 930 LEKKPNIDSRQEGCKPDCEGNELEFREVSEFYPCRPDVFILRGLSISERKTVAFVS 989  
 DB 1021 LERVQIDSDYDQGEKPKNCNSGVNVFQGVNPNYPTRPDIIVLQGLDISVXQGETLAVGS 1080  
 QY 990 SGCKSTSVOLLRVLPVQGVLPDGVDAKELNVQALRQOIAIVPEPLFNSTANI 1049  
 DB 1081 SGCKSTSVOLLRVLPVQGVLPDGVDAKELNVQALRQOIAIVPEPLFNSTANI 1140  
 QY 1050 AYGDNSRVPLDEIKERANANANHSIFTEGLPEKYNTOVGKGAOLSGGOKRLAIARAL 1109  
 DB 1141 AYGNMKNVYOELETKAKENANHSIFTESLQDKNTVGGKQRLSGGOKRLAIARAL 1200  
 QY 1110 QKPKILLDEATSLANDSEKVVQHALDKARTGTCVLVTHRSALONADLVYVHLGKI 1169  
 DB 1201 RKPKILLDEATSLANDSEKVVQHALDKARTGTCVLVTHRSALONADLVYVHLGKI 1260  
 QY 1170 KEQSTHOBELRNRIYFKLVNAQ 1192  
 DB 1261 VEOGTHOOLLQKGVYFSLVITQ 1283

RESULT 7  
 Q9T5U2 PRELIMINARY; PRT: 1163 AA.  
 AC Q9T5U2;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE MULTIDRUG RESISTANCE RELATED MRNA, PARTIAL CDS (FRAGMENT).  
 OS Felis silvestris catus (cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Okai Y., Nakamura N., Matsusiro H., Kato H., Setoguchi A., Yazawa M.,  
 RA Okuda M., Watarai T., Hasegawa A., Tsujimoto H.;  
 RT "Molecular analysis of multidrug resistance in lymphoma cells in the  
 cat.";  
 RL Am. J. Vet. Res. 0-0-0(1999).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS)  
 CC EMBL: AB029153; BA087071.1; -.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.

DR Pfam: PF00005; ABC\_tran: 2.  
 DR SMART: SM00382; AAA: 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER: 2.  
 DR ATP-binding: Transport.  
 KW NON\_TER 1  
 FT NON\_TER 1  
 SEQUENCE 1163 AA; 128510 MW; 1B5B413776A93A26 CRC64;

Query Match 55.2%; Score 3318.5; DB 6; Length 1163;  
 Best Local Similarity 55.2%; Pred. No. 1.5e-183;  
 Matches 644; Conservative 223; Mismatches 259; Indels 41; Gaps 8;

QY 11 NGACPLPPLVIGEMSDNL-----ISGLVOTNTYSPFRL-----TLVYV 50  
 DB 1 HGAALPLMLVFGMDTDSANANGSRNLTLLTNTGESIYNDYFNTRNLEEMTYTAYYS 60  
 QY 51 GIGVAAALIFGYIOISLMTITTAAROTKTRKOPFHSVLADODIGMPSDCDIGELNTRMT-DI 109  
 DB 61 GIGAGVLAAYIYQVFPCLLAAGROILKIRKOPFHAIMROEVGMFDYHDVDELNTRLTLDV 120  
 QY 110 DKISDGIDKIALLFQNNSTFSIGLAVGLVKGWKLTVLTSTPLIMASAAACSRWYISL 169  
 DB 121 SKINIGIGDKIGMFOQSMATFPIGFTVGRGWLTVLTLAISPYLGSAALWAKLISSEF 180  
 QY 170 TSKELSAVSKGAVAEVLSIRTYIAPRAQKELOQRTYONKDKADKDFIKRTIAKSVSL 229  
 DB 181 TDKELLAIAKAKGAVAEVLAIRTYIAPRAQKELEKRYKNNLEAKRIGIKAITANISII 240  
 QY 230 GAVYFMNGTYGLAWGYGSLTINCEPGYITGLVAVPFSVHSYSGICAAVPHETFAI 289  
 DB 241 GIAFLILTAAYALAWYGSIVLSHE--YSIGOVLFVFSVYLIGAFVQASPIEAFAN 298  
 QY 290 ARGAAHFHFOVLDKPSIDNFSYAGYKPSIEGTYEKNVSNFYSPRSIKILKGLNLRI 349  
 DB 299 ARGAAHYEIKIIDNKPISIDSYSGKNGKHPDNKGNLEKFNHVSYSRKEVKILKGLNLKY 358  
 QY 350 KSGEVALVGLNGSKSTYVQLOLRYPDDGFIMWDENDIALVNRHRIHGYVSOEP 409  
 DB 359 OSGQTVVALVGNSSGKSTVQLOLRYPDDGFIMWDENDIALVNRHRIHGYVSOEP 418  
 QY 410 VFEFTTISNNIKYGRDDVYDEEMERARANAYDFIMEFPNKNFTLYGKSGOMSGGQ 469  
 DB 419 VLEATTIENITKGRNNTMEIEKAVKANAYDFIMKPNKFDLTVGKQOLSGGQ 478  
 QY 470 RIARARALVNRPKILLDEATSAIDSESKASVQAALERAKSGRTTIVVAHRLSTIRSA 529  
 DB 479 RIARARALVNRPKILLDEATSAIDSESAVYQVALDKARKGRRTIVVAHRLSTIRNA 538  
 QY 530 IYTLDMGLAEKGAHAEIMAKRGLYSLVMSODIKKADOMESMTYSTERKTNLSPLHSV 589  
 DB 539 IAGFDGIVLVEKGNHDELMKEGIIYFKLVLMQ--TRKNLETELNAVYESISEIDALEM--SP 596  
 QY 590 KSIKSDFI-----DKAESSTOSKEISLPEVSLKILKILKPKMPPEVVLGT 634  
 DB 597 KSGSSLLRRKSTKRSIHASOGDKRLGTENKDERVPPVSWRILKILKILKMPPEVVLGT 656  
 QY 635 IASVINGTVHPVFSIIFAKITPMFGN-NDKTLKHDALDIYSLMIFVLIGYCEVSYFMQ 693  
 DB 657 FCATINNGIQAFAVSYLSRIIGVTRVEDPEPRKQNSIFSLLEFLVLGIISITPFLQ 716  
 QY 694 FYGRAGEITLMLRLHAFRAMLYODIAMPDEKENSGLTTLTLLADIDIOAGTSRIV 753  
 DB 717 TFGKAGEITLMLRLHAFRAMLYODIAMPDEKENSGLTTLTLLADIDIOAGTSRIV 776  
 QY 754 LFQONTNMGSLVYSPIFYEMETPILSLTAPLAVATGMETFAAMGFAKDKOELKHAK 813  
 DB 777 ITQNTANLSTGIITISLYWQTLTLLATITAIAGVEMKMLSGQALDKKKELEGAG 836  
 QY 814 IATELENITRTIVSLTRKAFQOMYEEMLOTOHRNTSKKAQIIGSCYAFSHAFYFAVA 873  
 DB 837 VATEAIENRTRYVSLTRKAFQOMYEEMLOTOHRNTSKKAQIIGSCYAFSHAFYFAVA 896

QY 874 GFRFGAYLIQAGRMTPEGMFIYFATALVAGMALIGTTLVLAPEYSKAKSGAHLFALLEK 933  
 DB 897 CRRFGAYLVAHEFPKDPQDVLVSAIVGMAVAGVSSAPDPAKAKYSAAHVIMIEKI 956  
 QY 934 PNIDRSQEGKKRDPCEGNLEFEVSEFFPCRPDPFILLGLSLIERGKTVAFVSSGGC 993  
 DB 957 PLIDSYSTEGMLPNLTLEGVTFNVEWFMNPTPRDIPVLOGLSLEVKKGOTLAVSSGGC 1016  
 QY 994 KTSVQLQRLYDPVQGOVLFDGVADKELNVQLSQAIVQPEVPLNCSIAENIAYGD 1053  
 DB 1017 KSTVQLLEFYDPMAGTVLIDGKEIKHLNVQMLRAHMGVISOEPILFDCSIGENIAYGD 1076  
 QY 1054 NSRVVPLDEIKRAANANHSFIEGLPEKYNVOYGLKGAOLSGGOKRATIRALLQRPK 1113  
 DB 1077 NSRVVQSEIIVAAKEANIHPIETLPDYNTRVGDGQTLSSGGOKRATIRALLQRPK 1136  
 QY 1114 ILLDEATSAIDNDESKVVOHALDKAR 1140  
 DB 1137 ILLDEATSAIDTESKVVQALDKAR 1163

RESULT 8  
 AC 088331 PRELIMINARY; PRT; 1321 AA.  
 AC 088331:  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE P-GLYCOPROTEIN SISTER PROTEIN.  
 GN SGP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN NCBI\_TaxID=10116;  
 RM [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Childs S.T., Yeh R.L., Hui D., Ling V.;  
 RT "Taxol resistance mediated by the liver-specific Sister gene of P-  
 glycoprotein";  
 RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS).  
 CC EMBL: AF010597; AAC24753.1; ...  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmam.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transport.  
 SEQUENCE 1321 AA; 146285 MW; 27F67EC36008D0F CRC64;

Query Match 47.2%; Score 2835; DB 11; Length 1321;  
 Best Local Similarity 45.4%; Pred. No. 1.8e-155;  
 Matches 571; Conservative 249; Mismatches 364; Indels 74; Gaps 10;

QY 1 MTLGLIASLVNCAQLPMLVIGEMSDNL-----SGLVOTNTY----- 40  
 DB 62 MLMGVCALLHMAQGLIIFGIMDPIKDYIERQELEIPGKACVNNVTIWINSSFHQ 121  
 QY 41 -----SFLRLTYVYGIGVAAALIFGYIOISLMTITTAAROTKTRKIRKOFHS 85  
 DB 122 NMTNGVGVADNIESLMKIFSGIYAGVMTVLLIGFQIRLWYITCARIRRMRTIYFR 181  
 QY 86 VLAQDGPWDSDDIGELNTRMT-DIDKISDGIDKIALLFQNNSTFSIGLAVGLVKGWKL 144  
 DB 182 IRMETGWDCHSVGELNSRFADIERINDAIADQLAHFLQHRSTFMCGLLIGFYRGWKL 241  
 QY 145 TLVTLSTPLIMASAAACSRWYISLTSKELSAVSKGAVAEVLSIRTYIAPRAQEKEL 204



```

Db 242 TLVILAVSPLIGIGAAVIGLSTIAKFTLELKAAYAKAGSIADVLSSIRTVAAFGENREV 301
QY 205 ORYTONLMDADDFGKRTIAKSVLSGAVYFPMNGTYGLAFWYGTSLINGEGYITIGVL 264
Db 302 ERYENLVPFAOMWGMWGMWGMFTGWTMCLFCYALAFWGSFLVID-EERYPGFLV 360
QY 265 AVFESVHSYICGAAYVHFFETAFARGAARHIEQVIDKPSIDMFSAQYKPESECTV 324
Db 361 QIFLCVILAAANNIGHASSCLEFSTGCSAATNIFOTIDROPYIDCMSDGYKIDIKKEI 420
QY 335 EFKVSENYPSRPSIKILKGLNLRIKSEYVALVGLNGSGSTVOLLQRLYDPDDGFM 384
Db 421 EFNHTFHPSPDPVKIILDNISMVIKPGETALVSGSAGKSTALQLORFEDPCEGMWT 480
QY 365 VDENDIRALNRYRHDHIGVYSGOEVLFGTTSSNNIKYGRDVTPEEMBARAENANADF 444
Db 481 LDGHDIRSLNRLWLDQIGIVEPVLSTTIAENIRFREDATMEDTVQAKKDANVNF 540
QY 445 IMEPNKNFTLVGEKGAQMSGOKORIAIARALVNPRLILDEATSAIDSEKSAYOVA 504
Db 541 IMALPQODFLVGEKGAQMSGOKORVAILARILNPKILLDMATSAIDNESEARVOEA 600
QY 505 LEKASKGTTIVVAHRISTINSADILVTLKQDMLAEKAHMLAKRLYSLVNSODI- 563
Db 601 LNKIOHGTIISVAHRLSTVRAADYIIGFEHGAVERGTHEELLERKGVYEMVLTLQSG 660
QY 564 -----KKADE-----QMSMTYSTERKINSI-----PLHSVKIK 593
Db 661 DNAHKETISIMKDATEGTELTERTSRGYSROSLRASIRKSSQSLTLTHDPPLAVADHK 720
QY 594 SDFIDKAEESTQSKESISPEVSLILKLNKPEMPFVLGTASVNLGTVHVFESIIPAK 653
Db 721 SSKNSKNDVLVEEVEEAPVR--RIIKYNIPEWHYIIIVGSSALINAVPIYLSLTSQ 778
QY 654 IITWFGNDKTTLKHDAIYSM--IFVLGVICVSYTMGGLFVGRAGEILTRLHLAF 711
Db 779 LIGTFSLIDKE--CORSEIHSKCLFEFVLIGCVSIPTQFLOGYTFKASELLTKRLKRGF 836
QY 712 KAMLYODIAMVEDKENSNGTLTILADIAIOGATGSSRIQVLTQNMNMGLSVIISLY 771
Db 837 KAMLGODIGMDDLRNNGVLTTRLATASVOGATGSOVGMVNSFNIIAALIAFF 896
QY 772 GWEMFTLLSIAPVLAVTGMETAMTGAFANKQKOLKHAGKATEALENRTIYSLTRE 831
Db 897 SMKLSLITIFEPFLALSGAVQTKMLTGFASDDQALEKAGITSEALSNIRTVAGIOVE 956
QY 832 KAFEBYEMQIOTQHRNTSKAQIIGSCYAFSAFIYAYAAFGFGAYLLOAGMPPEG 891
Db 957 GRTKAFEFVELQTSYKTVARRANITGLCFARFSQIAFLANSAAVRYGGYLLAEGIGFSH 1016
QY 892 MEIYETAAYGAMAIGKTLVLAPESKAKSGAHLFALLEKKNPNDSSOGKPKDPEG 951
Db 1017 VFRVYSSVVLATAVGRFFSTPTPAKAKISARFFOLLDRKRPINVSSEGEKMDNFOG 1076
QY 952 NLEFEVSEFFYPCRPDVEFLRGLSLISIRGKTVAEFGSSGCKSTSVOLLQRLYDPVOG 1011
Db 1077 KIDIFDKFTYPSRPDIQVNLGLISVNPQGTIAFVSSGGCKSISIDLERFYDPDGT 1136
QY 1012 VLFPGVDAKELNVMLSQIAIVPOEVLFLNCISIAENTAYGDNRSVPLDEIKENANAN 1071
Db 1137 VMIDGHSKKNIOFLRNSITIVSOEPLVFDSCISMDNITKYGDNTEIIVERAIAAKAQ 1196
QY 1072 IHSFEGLEPEYNTQVGLKAGNOLSGGOKORLAIARALLQKRIILLDDATSAIDNDSKV 1131
Db 1197 LHDFVMSPEPYETNVIGSQSLSGEKORIAIAIYVDRPILLLDDATSAIDPESKLT 1256
QY 1132 VOHALDKARTGTCIUVYHRLSAIONADLIYVHNGKIKEOGTHOELLRNRIYFKLV 1189
Db 1257 VQTALDKARBERTCIYVAHRLSTIONSDIIVAVSGGVIEKGTHEKMAOGAYIKLV 1314

```

```

045721
ID 045721 PRELIMINARY; PRT: 1294 AA.
AC 045721: 062101;
DT 01-JUN-1998 (TREMBLER, 06, Created)
DT 01-NOV-1998 (TREMBLER, 08, Last sequence update)
DT 01-JUN-2001 (TREMBLER, 17, Last annotation update)
DE C47A10.1 PROTEIN.
GN C47A10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RP SEQUENCE FROM N.A.
RA Basham V.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RM [2]
RX SEQUENCE FROM N.A.
RA MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproot J., Woulman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans."
RN Nature 368:32-38(1994).
RM [3]
RP SEQUENCE FROM N.A.
RA Basham V.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL: Z93782; CAB07855.1; -.
DR EMBL: Z81484; CAB07855.1; JOINED.
DR EMBL: Z81484; CAB03973.1; -.
DR EMBL: Z93782; CAB03973.1; JOINED.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR InterPro: IPR000130; Zn_fingerpds.
DR Pfam: PF00664; ABC_membrane_2.
DR SMART: SM00382; AAA; 2.
DR SMART: PS00211; ABC_TRANSPORTER; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR ATP-binding; Transport.
SQ SEQUENCE 1294 AA; 142490 MW; D8BED2602FA484DF CRC64;

Query Match 42.1%; Score 2529; DB 5; Length 1294;
Best Local Similarity 43.3%; Pred. No.9.3e-138;
Matches 533; Conservative 224; Mismatches 430; Indels 44; Gaps 10;

QY 1 MILGILASLVGACPLPPLVLVGENSDNLIS--GCLVOTN-----TYSFRL 45
Db 46 LAVGIVSCATGVGLPLMSIIMGNSQFVTLGTFIDPNSTASEKAAARAFESHVION 105
QY 46 TLVYVIGVVAALLFGYIOLISMITTAARQTRIRKQFHSVLAODIGKFDSCDIGELNTR 105
Db 106 CLKVYVLCGIFAGFLQASCFWVCEKLSNRFROFHSVMBEIAMVYDNTGTSLSNK 165
QY 106 MTD-IDKISDGIKIALLPQMSFFSGTGLAVGVKMKLTLYLTSPLMASAACS 164
Db 166 LFDNLERVRREGGVDVGLAFQMAQFVIGGFVAVLTYDMLTLTIMSSSPFMIGGLPLAK 225
QY 165 MVISLTSKELSAVSKAAVAEEVLSSIRTVIAFRAQEKELORYONLMDAKDFGIKRTIA 224

```



Db	226	LLAFAAKRAKQAVANQAGIAEVLUTSRVYANNGOBYECKRTEDALEHGKTKGKSF	285
Qy	225	SKVSLGAVYFFMNGTGYGLAFWGTSLLNGPGRYTGVLAVFYSHSSYIGAAPHF	284
Db	286	IGAGLAFSEFVITYASYCLAFWGTNFEYSGR--LESGVLTFVEFSVMGSMALGOAGOF	343
Qy	285	ETFAIAGAAPHFOYVDKRPISDINSTAGYKRESEIEGYEVRKNNFNPSPRSKILKG	344
Db	344	ATGTGALGAASLYEVIDRPEITDAYSTEBQRPISKISGRISVKNKEFYPTPRADKILKG	403
Qy	345	LNLRIKSGEVALVGLNGSGKSTFVOLLQRYDPDGFIMVENDIRALNVHYDHIYV	404
Db	404	VSLDAQGQYVALVNGSGCKSTIIDLQRYFPRDAQILDDIRFEDPNITKLYOLYGV	463
Qy	405	VSGPVLFGTTISNNIKYGRDQVYDEEMERAAEANAUDFIMEFPKFTULYGEKAOMS	464
Db	464	VSGPNTFNFTNISIQONITYGRSDVDEDIARALKEANAADFIKTFREGTLVIGDGYOMS	523
Qy	465	GGQORIALAIALYRNKILLLIDENASLDSEKSAVOALAEKASGRITTVARLSTI	524
Db	524	GGQORIALAIALYRNKILLLIDENASLDASESITVOSALENASGRITTVIARLSTI	583
Qy	525	RSADILYTLKQMLAEKGAHAELMARKGLYSLVMSO-----DIKKADQESMTYSYE	578
Db	584	RNAKTIYMKRAGOVMEVGTETLEIDQKGLYHELVAHQVADDDKPKKKEARRMSROTS	643
Qy	579	RKTNSLPLSHVSKISDPIDK-----AEBSTOSKEISLPE-----VSLKLTKILNK	624
Db	644	QRKGSV---NFKTQESQVDEKPGCAPARPEAKAEIKRIKLKEEBEBAVKANFLKLRRAR	700
Qy	625	PEPFVYVGLTASVLANCTVHPVESIIFAKITTFMGNDKTYTLKHDAEIVSMIFVLGYIC	684
Db	701	PEWYIFFEAIIALAIOGAVPAFSEFSQIITNFNSPDDOKMDHFMALFVILAQV	760
Qy	685	FVSYFMQGLEFGRAGELTYMLRNLHNLAKMALYODIAMPREDEKENSQGLTTLAIDLOIQ	744
Db	761	GTSMLFOCSLGVAAEKULMIRISKVYRANLRQDAIYFPMKHSIPRITTYRATLADAPINK	820
Qy	745	GATGSRIGQVLTONATNNGLSVSIISFYIGWMPFLILSIAPVLATVGMIFETAMTFPANKD	804
Db	821	SADIRYRGISTFNALASVGGGLGIAFYGCMARLVMAIFRPAAVGCALAMKKNHGSATSD	880
Qy	805	KOELHAKGKATLEALENIRITVSLGREKAEQOMEYEMLOTQHRNYSKKAQIIGSCYAFSH	864
Db	881	AKEMENAGKTAMEAIENIRFVOALTLQTKLYNIFCSHLDAPHGNSKAIIRGLTYGFAN	940
Qy	865	AFIFFAAAGRFCAVLI--QAGRMPREGFIYFATAIAGMAIGTIVLAPEYSKASG	922
Db	941	SIQFTYTAARREGLFLIFDKNVLMEBENVLRLVLAISFSFGTIGPMASYREPYIKAFPA	1000
Qy	923	AAHFALLEKKPNIDRSQEGKKPRDCEGNLEBEREVSFPYPCAPDVFILRGSLSTIEGK	982
Db	1001	AGLFNMLNEEPRRIDLGMTSSGTYPQ--LSGEVULKANKFFKFRPERPAVPIILOGIANVHKQO	1055
Qy	983	TVAFVSGSGCKSTSVOLLQRLYDPVOGOVLEFDGVDAKELANQWILRSQAIYPOEVLFN	1042
Db	1060	TLALVPSGCGKSTVYSLBRLYDPLEGAVYDNDMLRQMNKHLKHTIALVSOEPILED	1119
Qy	1043	CSIAENIATYGDNSVWVLDIEKEAANANIHSTBELPERKYNTQYGLKGAOLSGQOKORL	1102
Db	1120	TSIENIIVYGIPOPEYTHQHEIFACSNANIKHKEIDELPGUYETRVGEKGTQOLSGQOKORI	1179
Qy	1103	AIAFALLQKPIILLDBATSAIJDNSKVVQVQHALLDKARGICIVYTHLSIAIOWADLIV	1166
Db	1180	AIAFALLNPFILLLIDENATISALDTESKVOVQVALLDAADKRTCIYVAHRLSTIVAGCTM	1239
Qy	1163	VLANGKIREOGTHOELLRNSDIFKVLNAOS 1193	
Db	1240	VVKNGVQVVEGTHNELAKRGATFALLOKOS 1270	
RESULT	10		

09W693			
AD	Q9W693	PRELIMINARY;	PRT; 851 AA.
IC	Q9W693;		
DT	01-NOV-1999 (TREMBLrel, 12, Created)		
DT	01-NOV-1999 (TREMBLrel, 12, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)		
DE	MULTIDrug RESISTANCE TRANSPORTER HOMOLOG (FRAGMENT).		
GN	MDR.		
OS	Fundulus heteroclitus (Killifish) (Mummichog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei		
OC	Acanthomorpha; Acanthopterygii; Perciforma; Atherinomorpha;		
OC	Cyprinodontiformes; Fundulidae; Fundulus.		
OX	NCBI_TaxID=8078;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=LIVER, INTESTINE;		
RA	Cooper P.S., Van Veld P.A., Reece K.S.;		
RT	"P-glycoprotein related sequences from the mummichog (Fundulus		
RL	heteroclitus)".		
CC	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABCG		
CC	TRANSPORTERS).		
DR	EMBL; AF099732; AA02356.1; -.		
DR	InterPro; IPR003593; AAA.		
DR	InterPro; IPR001140; ABC_transporter_tmam.		
DR	InterPro; IPR003439; ABC_transporter.		
DR	InterPro; IPR001687; ATP_GTP_A.		
DR	Pfam; PF00064; ABC_membrane_1.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	SMART; SM00382; AAA; 1.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
DR	ATP-binding; Transport.		
FT	NON_TER		
FT	1		
SO	SEQUENCE	851 AA; 93660 MW; 8F2DC4464171420A CRC64;	

[illegible]



Db 879 LVMGVEFMQMRFGKQIDTDLLEAGKVASQAVEHIRTVSLHNQEOGHFTYCEVLRP 938  
Qy 846 HRNTSKAQIIIGSCYAFSHAFIYFAYAGFRGATLIIQAGRTPEGMFVFAIAYGAMA 905  
Db 939 FNTNLKHAHTGAVAFQSOLFMYAAFYIGSIFVNOQAQPIIDYRVFPAISFCGM 998  
Qy 906 IGTTLVLAPEYKSAKSGAAHFALEKKNPINDRSOEG-KKPDCEGNELEFEVSPYPC 964  
Db 999 IGTNTSFIPDYVKARLALSLFLYLEHPTPIDSLSGIVKRT--GNISINNVFENPT 1056  
Qy 965 RPDVEILGLSISIRGTVAFVSGGKSTVOLLRLYDPVOGOVLFDGVDKELNV 1024  
Db 1057 RKDTVLGGLFDLIRAKGTVALVGHSGGKSTIMGLERFYHQDKMIDGNIRNLNI 1116  
Qy 1025 OMLRQIAIVQEPVLFNCSTIAENIAYGDSRVPLDEKKAANANHSFIEGLEPEKN 1084  
Db 1117 SFLRQVQVISOEPTLFCCTIGENICVGTN-RNVYIOEIVEAKAMANHFTLIGLPGYD 1175  
Qy 1085 TVGGLKGAOLSGGKORLAIRALLQPKILLDEATSAIANDSEKVVQHALDKARTGT 1144  
Db 1176 TVHGKGTQLSGGKORLAIRALVSPSVLLDEATSAIDSEKIVEOALDAKQGT 1235  
Qy 1145 CLVYHRLSAIONADLIYVHLNGKIKEQTHQELRNRIYKLVNAOSV 1194  
Db 1236 CLVIAHRLSTIONSDVIAVSGKIVEKTHDELIRKSEITVQFECTORI 1285

RESULT 12  
Qy 09V626 PRELIMINARY; PRF: 1313 AA.  
AC 09V626  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE MDR50 PROTEIN.  
GN MDR50 OR CG8523.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta.  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Dudkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
Hoskins R.A., Houston K.A., Howland T.J., Wei M.-H., Ibeigwan C.,  
Jostali D., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
Palatzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Relneert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster.";  
Science 287:2185-2195(2000).  
RL -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
CC TRANSPORTERS).  
DR EMBL: AE003815; AAF58271.1; -.  
DR Flybase: FBgn0010241; Mdr50.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane\_2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR ATP-binding; Transport.  
KW SEQUENCE 1313 AA; 145277 MW; 3B99DEA690AAB58 CRC64;

Query Match 39.7%; Score 2389; DB 5; Length 1313;  
Best Local Similarity 40.8%; Pred. No. 1.2e-129;  
Matches 510; Conservative 233; Mismatches 422; Indels 84; Gaps 16;

Qy 2 IIGTILASLVNGACLPMPVILGEMSDNLT--SGCLVOTMTY-----SFRRLT 47  
Db 89 VIGLLSAVATGTLTPANSLSIFGLNLAIDGLLESKGSYRADAISTLLDKYQFSL 148  
Qy 48 YYVGIVGVALIFGYIOISMITTAAROTRRIRKQFHSVALDIGNFDSCDIGELTMRM 107  
Db 149 QMTYIGIMLVCSYSTSTFCFNAAHSQILITRSKFRSLIHQDMKYDNDNOSEVASRN 208  
Qy 108 -DIDKISDGIKIALLFQNMSTFSTIGLAVGVKMKLTLYLSTSPILMAS---AAAC 162  
Db 209 EDLSKMEIDLAKRVVVFVYLVAVFVGLVAFVGMQLSLVCLTSLFLFIAMGLVAVAT 268  
Qy 163 SRMVISLPTSKELSNYSKAAVAEEVLSSTRYIAFAQKELEQRYONLKDANKDFIKRT 222  
Db 269 SR---LAKREYTMAGAAVAVAGALSGIRYKAFFEGEAKVAAYKERVAAKILINRN 324  
Qy 223 IASKVSLGAVFFFMNGTYGLAFWYGTSLILNG--EP---GYTIGYLVAFVSVYHSSYCI 277  
Db 325 MSGIGFGLMFFIYASVYALWYGVGLYIKGHEPAYENYDAGIMTYFFSVMGSMNI 384  
Qy 278 GAAPHFETFAIARGAFAHIFOVIDKRPISIDNESTAGYKPESTIEGVEKNSFNPSPR 337  
Db 385 GMAAPYIEAFGLAKGACAKVAFHIEQIPEINPIDGSKKLNELTITIEFEVEFYQPTRP 444  
Qy 338 SIKILKGLMLRKSGTVALVGLNGSKSTVQOLLRLDPPDGFIMDENDIRALNVH 397  
Db 445 EVSLINKMLKILHRGTVALVGVSGGKSTCIOLVRFYDPOAGNLLFGNTLNKLDIWM 504  
Qy 398 YRDHGVGQSEPELFTFTTNSNIKYGRDVTDEEMRARERANAYFIMEFPKFTVLV 457  
Db 505 LMSRIGVQOEPLFTSTYENIRYGRDATREIEIAAAAANAALFIKKLPGRGYTLVG 564  
Qy 458 EKGAGMSGOKORAIARALVRNPKILLDEATSAIDSEKSAVAQALERKSGRTTVV 517  
Db 565 EGAQGLSGGOKORAIARALIRDPETILLDEATSAIDTASEAAVQALERKVSAGRTTV 624  
Qy 518 AHRLTSTRADLIYTLKDKMAEKGAHAELMAKRGILYGVSMQ-----561  
Db 625 AHRLTSTRADRIYLVNKEVEESGTHQELMELKDYFNLVTTQGLGEDGSLVSPGDIY 684  
Qy 562 ---DIKKADEOMESMTYSTERKTNISLPLHSVKSISKD-----FIDKAESTQSKISLP 612

Db 685 KNFDKDEDEE-----EIKVSEDEDEDVNVTEKNNKKKKKKVDP 726  
 QY 613 -EVS-LTKILKLNKPEMPFVVLGLTASVLTGVHPVPSIIFAKITTFMG--NNDKTLKH 668  
 Db 727 NEKPMLEEVKMKPELQIIVAGCISVIMGCAMPFAVLEFGLTQILSKDNOQ-YVRE 785  
 QY 669 DAEIYSIMFVILGVICVSYFMQLEFYGRAGEIITMKLRHLAKMLYODIAMFERKNS 728  
 Db 786 NSNOYSLYFIAGIIVGIAITFLOIYFFGIAGERLTERLGRIMFEAMLRQEVAMPDDKANG 845  
 QY 729 TGGTITLLADIDAOIGATSGRIGVLQONATNMGLSVIISFYGEMTFLLISAPVLAV 788  
 Db 846 TGSICARLSDDAAVQATGORTTIVQISITLALGIALSYEMSGVLVALATPFLI 905  
 QY 789 TGMETATMTGFANKDQKOEKHAQKITEALENIRITVLSITREKAFOMYEML-----QT 844  
 Db 906 AFTWQRTLMKKNMGSKKTMYENCKLAVEVYSNRTVYASLGRMEFMQNTIGMLIPAVEI 965  
 QY 845 QHRTSKKAOIISCVAFSHAFYFAYAGFRGAYLLIQGRMTPEGMFTVETAAYGAM 904  
 Db 966 SKRTHRR-----GLVGLASLMEFAYAACMYGTWCIVHGGILFGDVFKVSAQLMGTA 1021  
 QY 905 AIGTIVLAIEYKSAKGAHLPALLEKPRIDRSQBGKPPDCEGNLEFREVSEFYPC 964  
 Db 1022 STANALAFAPNMOKGVAATITFELRQPSYDRPGVSRDPMHSEGYVREDKVKFSYPT 1081  
 QY 965 RPVEILRGLSLSTERKTVAFVSGSGCKSTVOLRLQRLYDPVOGOVLFPDGVAKELNV 1024  
 Db 1082 RSEIYVKGLELAVSKQKIALVGPSCGKSTCQLQRFYDVEGATLIDECVAVNSM 1141  
 QY 1025 QWRSQAIYVQEPVLENCSTAIENIAYGDSRVVPLDEIKEAANAANIHSFIEGLPEKYN 1084  
 Db 1142 TNLNQGIVSQEPIILDRITRENISYGDNARNYTDDEIISACKSNHETIANMLPLGYD 1201  
 QY 1085 TQVGLKAOISGGOKRLATARALLOKPIILLDEATSAIDNSEKVVQADLAKRGT 1144  
 Db 1202 TRMEKKAQISGGOKRIARALIRNPKIMLDEATSAIDSEKVVQADLAKRGT 1261  
 QY 1145 CLVTHRLSAIONADILVHLNGKIKEQGTQHOELLRNNDIYFKLVNQS 1193  
 Db 1262 TISTAHRLSTVYHSDVLFVFENGVLVCEAGDHKQLANGLYTYLTKIQS 1310  
 RESULT 13  
 ID 061301 PRELIMINARY; PRT: 1275 AA.  
 AC 061301:  
 DT 01-AUG-1998 (TEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
 DE P-GLYCOPROTEIN.  
 GN PGP-A.  
 OS Haemonchus contortus.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylidae;  
 OC Trichostrongylidae; Haemonchidae; Haemonchinae; Haemonchus.  
 OC NCBI\_TaxID=6289;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98226176; Pubmed=9566525.  
 RA Xu M., Molento M., Blackhall W., Ribeiro P., Beech R., Prichard R.;  
 RT "Ivermectin resistance in nematodes may be caused by alteration of p-  
 glycoprotein homolog";  
 RL Mol. Biochem. Parasitol. 91:327-335(1998).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 CC EMBL; AF003908; AAC38987.1; -;  
 DR HSSP; P13569; INBD.  
 DR InterPro; IPR003593; AAA.  
 DR InterPro; IPR001140; ABC\_transporter\_tmam.  
 DR InterPro; IPR003439; ABC\_transport.  
 DR InterPro; IPR001687; ATP\_GTP\_A.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.

DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW ATP-binding; Transport.  
 SO SEQUENCE 1275 AA; 142203 MW; 7439D3836B75787C CRC64;  
 Query Match 39.6%; Score 2380.5; DB 5; Length 1275;  
 Best Local Similarity 40.9%; Pred. No. 3,6e-129;  
 Matches 502; Conservative 244; Mismatches 431; Indels 51; Gaps 13;  
 QY 4 GILASLVNGACLPMLPVLGEM-----SDNLSGCL-----YQTNYSF-----R 44  
 Db 54 GTLAPCVHGAAGFVSGLVGLGMMTVFLRAQNSEFVLQTSRDEGDLALKEEDTVLR 113  
 QY 45 LTIYVIGVVALTFGYTQISLMTITTAAROTKRIRKOFHFSVLAQDIGMFDSCIGELANT 104  
 Db 114 YCLTYLGLGFAMFATSYIQIYCNHETFAERITHKRIYLAAILRQISWPDIOQTGLTLA 173  
 QY 105 RMT-DIDKISDGIDKIALLEFQNNSTFSGIANGLVKWKILVTLTSTPLIMASAAACS 163  
 Db 174 RLTDLERVRGLDKLSLFIQWYSAFVAGFCVGFAYSMWSMTLVMAVVADEFIVISANMMS 233  
 QY 164 RMVLSLSKELSAKAGAAVEVLSIRIVYARAEKELQRTQNLKDAKDGIRTI 223  
 Db 234 KIVATRTQVDEETVAVGALAEFTFSSIRTVHSICGHRRELTREALDEKRGOTGLVKYF 293  
 QY 224 ASKYSIAVYFFMNGTGLAFWYGTSLINGEPGYTIGTVLAVFYSYHSYICGAAPVH 283  
 Db 294 YMGVGVFGCMCTVYSALAFWYGSVLIN-DPALDGRIFTVFVFMSSAALGTCLPH 352  
 QY 284 FETFAIRGAAGAFHIFQYIDKPKSIDNFSTAGYKRESIEGVFEFNVSEFNPSPRSITLK 343  
 Db 353 LNTISIRGAARVSLVSYNSRPKIDPYSLDGIIVNNRGSIRFNHVSYPSPRRTDILK 412  
 QY 344 GLNIRISGEVVALVANGSGKSTVYOLRLYPPDDGFILVNDNRLALNRYRHIG 403  
 Db 413 GVSILQVSGKIALVSGSGCKSTNNMLRFLYPTGKATYIDDIQCDLNVQKLRQIG 472  
 QY 404 VVSQEPVLFGTTSNNIKYGRDVTDEMERARAEANAYPEIMEPNKENTLVGEKGAQM 463  
 Db 473 VVSQEPVLFQSTLEFNKMGVEQATMEVOACVANAADPTKRLPGYGRVGERVOL 532  
 QY 464 SGGOKRIARALVRNPKIILLDEATSAIDSEKSAVQAALAKASGRITTVAAHRLST 523  
 Db 533 SGGOKRIARAIKIRNRIILLDEATSAIDTEAEISYQEALEKAKGRITVVAHRLST 592  
 QY 524 IRSADILYTLKGMIAEKGAHAELMAKRGLYSLVMSODIK--KADOMESMTSTERT 581  
 Db 593 IRNVDQIFVFENKGTIVDQTHAEIMNKRGVFEEMTQAOVLKQKEEVEVDSDAES-DVVS 651  
 QY 582 NSLPLHSAVKSISPDFIDAEESTOSKEISLPEV-----SLKILKLNKPE 626  
 Db 652 PDIALPHLSIURS-----RKSTSRSAISAVPSVSMQEMEDLAKPTPKIKIIFYNRDK 706  
 QY 627 WPEVVLGLTASVLTGVHPVPSIIFAKITTFMGNDKTTLKHADEIYSIMFVILGVICFV 686  
 Db 707 WGYFLILGLIACITIGTVPTFAVLAYQIIQYV-SEPDQMKGHLVFCGAFIYIGLVHAF 765  
 QY 687 SYFMQLEFYGRAGEIITMKLRHLAKMLYODIAMFERKNSGGLTITLLADIDAOIGA 746  
 Db 766 AFFESAIQLGRGELTKLFEFAFKLLRONVGPYDRIHGTGKLCITRFATDAPNHYV 825  
 QY 747 TGSIRGLVLTQONATNMGLSVIISFYGEMTFLLISAPVLAVTGMETATMTGFANKDQ 806  
 Db 826 FTRLEPGLSVSYITIG-ALVIGLIFGQMLALILWVMPPLIGSGYFENRMQFGKMRDTE 884  
 QY 807 ELKHAGKATEALENIRITVLSITREKAFOMYEMLQTOHNTSKKAOIISCVAFSHAF 866  
 Db 885 LLEBAGKVAQAVERNIRFVALNRQEOHFHMYCEVLEKPYRENLCQAHYTGVAFAFSQSL 944  
 QY 867 IYFAYVAGFRGAYLLIQGRMTPEGMFTVETAAYGAMAIKTLVLAPEYKSAKGAHL 926  
 Db 945 LFMVAVAFWIGALFVDNHSQPIIDVYRVFPAFEMFCQGMVGNISSFTPDVVKARLASL 1004

[illegible]

RESULT 14  
080725  
ID 080725 PRELIMINARY;  
PRT; 1286 AA

DT 01-NOV-1998 (Tremblrel, 08, Created)  
DT 01-NOV-1998 (Tremblrel, 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel, 17, Last annotation update)  
DE PUTATIVE P-GLYCOPROTEIN.  
GN F14M4.17.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Euryalotia: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxId=3702;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,  
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;  
RT "Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence."  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -I- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
CC TRANSPORTERS).  
DR EMBL; AC004411; AAC34225.1; -.  
DR HSSP; P13569; 1MBD.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transporttr.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR InterPro: IPR001525; C5\_DNA\_meth.  
DR Pfam; PF00664; ABC\_membrane; 2.  
DR Pfam; PF00065; ABC\_tran; 2.  
DR SMART; SMO0382; AAA; 2.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
DR PROSITE; PS00095; C5\_MTASE\_2; UNKNOWN\_1.  
DR ATP-binding; Transport.  
SQ SEQUENCE 1286 AA; 139027 MW; 97D2A8AFBA698E6 CRC64;

Query Match	39.1%	Score 2351.5	DB 10	Length 1286;
Best Local Similarity	40.5%	Pred. No. 1.7e-127;		
Matches 496;	Conservative 233;	Mismatches 430;	Indels 47;	Gaps 14

```

Oy 1 M1GLTSLVNGACPIPLPVLVGGSSNNLSGCLVQNTNY--SEFRLLLYLVGIGVALIE 59
Db 63 M1GLTSLGSGNGCLSPPLMTLLTFGDLIDAFGEN---QNTNTDKYSKALKPFWLIGIGFEEA 119
Oy 60 GYIQLSLMTITTAQTKRIKQKFHSVLAQDIGCFD--SCDIGELNTNTMT-DIDKISDGI 117
Db 120 AFLQISLGMWIGGEQQAARISRLYKTLTIRODAIFDLIDNTNGEVVNRSGVLYIDPAMG 179

```

[illegible]

Db 1253 KNGVIEKGMHDTLINIKDGVASLV 1278

RESULT 15

ID 014573 PRELIMINARY; PRT; 469 AA.  
AC 014573;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE WUGSC:H\_RG367017.1 PROTEIN (FRAGMENT).  
GN WUGSC:H\_RG367017.1  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kallio J., Smith A., Gibson A.;  
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Waterston R.;  
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).  
DR EMBL: AC002486; AAB67603.1; -  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane.1.  
DR Pfam: PF00005; ABC\_tran.1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER.1.  
KW ATP-binding; Transport.  
FT NON\_TER 1 469  
FT 1  
SQ SEQUENCE 469 AA; 51667 MW; E67E088FF594FC3 CRC64;

Query Match 39.08; Score 2346.5; DB 4; Length 469;  
Best Local Similarity 99.84; Pred. No. 6,4e-128;  
Matches 468; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 40 YSFRLFLYYGIGVALIFGIQISLWITFAROTFRIRKOPFHSVLAODIGWFSDDI 99  
DB 1 YSFRLFLYYGIGVALIFGIQISLWITFAROTFRIRKOPFHSVLAODIGWFSDDI 60  
QY 100 GELNTRMT-DIDKISDGIKIALLFQNMSTFSGLAVALGKMKLTLVLTSLTSLIMAS 158  
DB 61 GELNTRMTDIDKISDGIKIALLFQNMSTFSGLAVALGKMKLTLVLTSLTSLIMAS 120  
QY 159 AAACSRVVISLTSKELSAISKAGAAEVLSSITVIAFRAQEKELQRYTONLKDADFG 218  
DB 121 AAACSRVVISLTSKELSAISKAGAAEVLSSITVIAFRAQEKELQRYTONLKDADFG 180  
QY 219 IKRTIAKSVISGAVYFPMNGYGLAFWGTGLIINGEPGYTGVLAFVFSVHSSYICG 278  
DB 181 IKRTIASVSVISGAVYFPMNGYGLAFWGTGLIINGEPGYTGVLAFVFSVHSSYICG 240  
QY 279 AAVPFETFAIARGAAPFIQVIDIKKPSIDNFSTAGYKPEISIEGTVERRKNSFNYPSPRS 338  
DB 241 AAVPFETFAIARGAAPFIQVIDIKKPSIDNFSTAGYKPEISIEGTVERRKNSFNYPSPRS 300  
QY 339 IKILKGLNLRIKSGETVALVGLNGSGKSTVOLLQRLYDPDDGFLMVDENDIRALNVRHY 398  
DB 301 IKILKGLNLRIKSGETVALVGLNGSGKSTVOLLQRLYDPDDGFLMVDENDIRALNVRHY 360  
QY 399 RDHIGVVSQEPVLFGTTSNNIKYGRDDVTDDEMRARAEMANAYDFIMEFPNKFTLVGE 458  
DB 361 RDHIGVVSQEPVLFGTTSNNIKYGRDDVTDDEMRARAEMANAYDFIMEFPNKFTLVGE 420  
QY 459 KGAOMSGGOKORIAIARALVNRPKLILDEATSAIDSESKSAVOAALEK 507

Db 421 KGAOMSGGOKORIAIARALVNRPKLILDEATSAIDSESKSAVOAALEK 469

Search completed: April 22, 2002, 14:37:04  
Job time: 256 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:37:38 ; Search time 24.77 Seconds

(without alignments)  
800.794 Million cell updates/sec

Title: US-09-873-409-7

Sequence: 1 MILGILASLVNACLPMLP.....SEKSAVQALEKDPFRYSF 541

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1566.5	57.2	1276	1	MDR1_MOUSE
2	1555.5	56.8	1280	1	MDR1_HUMAN
3	1555	56.8	1276	1	MDR2_CRIGR
4	1548	56.5	1277	1	MDR1_RAT
5	1542	56.3	1276	1	MDR3_MOUSE
6	1541.5	56.3	1276	1	MDR3_CRIGR
7	1455.5	53.1	1279	1	MDR3_HUMAN
8	1440.5	52.6	1281	1	MDR3_CRIGR
9	1439.5	52.6	1276	1	MDR2_MOUSE
10	1421.5	51.9	1278	1	MDR2_RAT
11	1250	45.6	1321	1	AB11_HUMAN
12	1220	44.5	1321	1	AB11_RAT
13	1204	44.0	1321	1	AB11_MOUSE
14	1153	42.1	1321	1	AB11_RABIT
15	1002	36.6	1302	1	MDR4_DROME
16	997.5	36.4	1321	1	MDR1_CAEEL
17	949	34.6	1234	1	MDR3_CAEEL
18	925	33.8	1362	1	PMDL1_SCHPO
19	911.5	33.3	1280	1	MDR1_LETEN
20	908.5	33.2	1302	1	MDR5_DROME
21	676.5	24.7	735	1	ABCB_HUMAN
22	673	24.6	820	1	MDL2_YEAST
23	669.5	24.4	1419	1	MDR5_PLAFR
24	646	23.6	695	1	MDL1_YEAST
25	640	23.4	685	1	MDL1_CANAL
26	638	23.3	703	1	TAP2_RAT
27	626.5	22.9	686	1	TAP2_HUMAN
28	622	22.7	702	1	TAP2_MOUSE
29	622	22.7	726	1	YFX9_SCHPO
30	598.5	21.9	724	1	TAP1_MOUSE
31	596	21.8	725	1	TAP1_RAT
32	582.5	21.3	582	1	MSBA_ECOLI
33	575.5	21.0	598	1	Y288_THEMEA

34	574.5	21.0	748	1	TAP1_HUMAN	O03518 homo sapien
35	557.5	20.4	587	1	MSBA_HAEIN	P44407 haemophilus
36	551.5	20.1	604	1	YFIC_BACSU	P54719 bacillus su
37	536	19.6	532	1	HEPA_ANASP	P12638 anabeena sp
38	533	19.5	1290	1	STE6_YEAST	P12866 saccharomyc
39	529	19.3	586	1	YAGM_RHISN	P55469 rhizobium s
40	528	19.3	1743	1	TAGC_DICDI	O23868 dictyostell
41	513	18.7	1589	1	Y015_MYCCE	P47261 mycoplasma
42	504.5	18.4	707	1	HLX2_ECOLI	P10089 escherichia
43	501.5	18.3	707	1	HLXB_ECOLI	P08716 escherichia
44	500.5	18.3	707	1	HLXB_PROVU	P11599 proteus vul
45	496	18.1	707	1	RT1B_ACPPL	P26760 actinobacil

## ALIGNMENTS

RESULT	ID	MDR1_MOUSE	STANDARD:	PRT: 1276 AA.
AC	P06795:			
DT	01-JAN-1988	(Rel. 06, Created)		
DT	01-JAN-1988	(Rel. 06, Last sequence update)		
DT	20-AUG-2001	(Rel. 40, Last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).			
GN	ABCB1 OR PGY1 OR PGY1-1 OR MDR1 OR MDR1B.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87028229; PubMed=3768958;			
RA	Gros P., Croop J., Housman D.;			
RT	"Mammalian multidrug resistance gene: complete cDNA sequence			
RT	indicates strong homology to bacterial transport proteins.";			
RL	Cell 47:371-380(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89367274; PubMed=2570420;			
RA	Raymond M., Gros P.;			
RT	"Mammalian multidrug-resistance gene: correlation of exon			
RT	organization with structural domains and duplication of an ancestral			
RL	gene.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).			
RN	[3]			
RP	SEQUENCE OF 1-21 FROM N.A.			
RX	MEDLINE=91042535; PubMed=2248681;			
RA	Raymond M., Gros P.;			
RT	"Cell-specific activity of cis-acting regulatory elements in the			
RT	promoter of the mouse multidrug resistance gene mdr1.";			
RL	Mol. Cell. Biol. 10:6036-6040(1990).			
CC	-1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED			
CC	DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- PTR: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE			
CC	LINKER DOMAIN.			
CC	-1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE			
CC	RELATED BUT DISTINCT CELLULAR GENES.			
CC	-1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY			
CC	(ABC TRANSPORTERS). MDR SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M14757; AAA79005.1; -			
DR	EMBL; M60348; AAA39513.1; -			
DR	PIR; A33719; DVMS1.			

DR MGD; MG1:97568; Abcb1.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001140; ABC\_transport\_tmam.  
 DR InterPro: IPR001687; ATP-GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
 KM Multigene family; Phosphorylation.  
 FT DOMAIN 1 47  
 FT TRANSMEM 48 71 POTENTIAL.  
 FT TRANSMEM 119 139 POTENTIAL.  
 FT TRANSMEM 199 219 POTENTIAL.  
 FT TRANSMEM 296 316 POTENTIAL.  
 FT TRANSMEM 327 347 POTENTIAL.  
 FT DOMAIN 348 709 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 710 730 POTENTIAL.  
 FT TRANSMEM 755 775 POTENTIAL.  
 FT TRANSMEM 831 851 POTENTIAL.  
 FT TRANSMEM 855 882 POTENTIAL.  
 FT TRANSMEM 935 955 POTENTIAL.  
 FT TRANSMEM 966 986 POTENTIAL.  
 FT DOMAIN 987 1276 CYTOPLASMIC (POTENTIAL).  
 FT NF\_BIND 426 433 ATP (BY SIMILARITY).  
 FT NF\_BIND 1068 1075 ATP (BY SIMILARITY).  
 FT REPEAT 1 635  
 FT REPEAT 636 1276  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 1276 AA; 140993 MW; 1804D0F011B0F4E CRC64;

Query Match 57.2%; Score 1566.5; DB 1; Length 1276;  
 Best Local Similarity 54.9%; Pred. No. 1.9e-100;  
 Matches 307; Conservative 99; Mismatches 94; Indels 59; Gaps 6;  
 QY 1 MITGLTSLVNGACLPPLVPLVLEKMSDNL-----SCCLVQNT----- 39  
 DB 50 MITGLTSLVNGACLPPLVPLVLEKMSDNL-----SCCLVQNT----- 39  
 QY 40 ---VSFRLTYVYIGIAALIFGYIOISLITIAAROTKRIRKOFPSVLAADIGWDS 96  
 DB 110 MATAVY-----YTGIGAGVLIYVIOVSLWCLAGKQIKIRKOFPSVLAADIGWDS 164  
 QY 97 CDIGELNTRMT-DIDKISDGIKRIALLFQNMSTFSGIAGLVKGMKLTVLSTPLI 155  
 DB 165 HDVGLNTRMT-DIDKISDGIKRIALLFQNMSTFSGIAGLVKGMKLTVLSTPLI 155  
 QY 156 MASAAACSRWVISTLSKELSYKAGAAVEVLSIRIVIAFRAOEKELQSFLLNTRY 215  
 DB 225 GLSSALMAKVLSTFNKELQAYAKAGAAVEVLSIRIVIAFRAOEKELQSFLLNTRY 215  
 QY 216 AMEYFPQWLSCVLFKPYATQNLMDKADGFKRTIASVSGAYVFENNGVYGLAFWGT 275  
 DB 275 -----RYNNLEAKNVGIRKATIASIGIALLVYASTALAFWGT 317  
 QY 276 SLINGEGTYTIGVLAFFSVIHSYICGAAPHFETFAARGAFAFIQVDDKPSID 335  
 DB 318 SLVLSNE--YSIGEVLYTFSSILGTFSIGHLARIEFAANARGAFFIRIINDEPSID 375  
 QY 336 NPSAGYKPEISIECTVEKFNPNPSPRSKILKGLNLRIKSGFTVALVGLNSGKSTV 395  
 DB 376 SFSYKGYKPDISIMNLEKFNPNPSPRSKILKGLNLRIKSGFTVALVGLNSGKSTV 395  
 QY 396 VOLLORLYDPPDGFIMVDENDIRALNVHYHDHGVSGEVLPTTISNNIKYRDVY 455  
 DB 436 VOLLORLYDPPDGFIMVDENDIRALNVHYHDHGVSGEVLPTTISNNIKYRDVY 455  
 QY 456 DEEMERARERANAYDFIMEFNKFTLVGEGAGMSGGOKRIATARALVRNPKILIDE 515

DB 496 MDEIEKAVKEANAYDFIMKLPHPQFDTLVGERQALSGGOKRIATARALVRNPKILIDE 555  
 QY 516 ATSLADESKSAVOALEK 534  
 DB 556 ATSLADESKSAVOALDK 574  
 RESULT 2  
 MDRL\_HUMAN STANDARD; PRT; 1280 AA.  
 ID MDRL\_HUMAN  
 AC P08185; Q12755; Q14812;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).  
 GN ABCB1 OR PGY1 OR MDRL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RX MEDLINE=87028230; PubMed=2876781;  
 RA Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,  
 "Internal duplication and homology with bacterial transport proteins  
 in the mdrl (P-glycoprotein) gene from multidrug-resistant human  
 cells.";  
 RL Cell 47:381-389(1986).  
 RN [2]  
 RX MEDLINE=90094448; PubMed=1967175;  
 RA Chen C.-J., Clark D.P., Ueda K., Pastan I., Gottesman M.M.,  
 "Genomic organization of the human multidrug resistance (MDR1) gene  
 and origin of P-glycoproteins.";  
 RL J. Biol. Chem. 265:506-514(1990).  
 RN [3]  
 RX MEDLINE=97190336; PubMed=9038218;  
 RA Chen G., Duran G.E., Steger K.A., Lacayo N.J., Jaffrezou J.P.,  
 Dumoncel C., Slikic B.I.;  
 RT "Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein,  
 altered phenotype, and resistance to cyclosporins.";  
 RL J. Biol. Chem. 272:5974-5982(1997).  
 RN [4]  
 RP SEQUENCE OF 1-234 FROM N.A.  
 RA Smith A., Beck C., Gibson A.;  
 RL submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 178-215 AND 800-856 FROM N.A.  
 RA MEDLINE=90290529; PubMed=1972623;  
 RA Geleler V., Weger S., Probst H.;  
 RT "mdrl1/P-glycoprotein gene segments analyzed from various human  
 leukemic cell lines exhibiting different multidrug resistance  
 profiles.";  
 RL Biochem. Biophys. Res. Commun. 169:796-802(1990).  
 RN [6]  
 RP SEQUENCE OF 1-23 FROM N.A.  
 RA Klocka N., Tsubota J., Kakehi Y., Komano T., Gottesman M.M.,  
 Pastan I., Ueda K.;  
 RL submitted (JUL-1991) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SMALL INTESTINE  
 CC AND BRAIN.  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M14758; AAA59575.1; -  
DR EMBL: M29447; AAA59576.1; -  
DR EMBL: M29424; AAA59576.1; JOINED.  
DR EMBL: M29425; AAA59576.1; JOINED.  
DR EMBL: M29426; AAA59576.1; JOINED.  
DR EMBL: M29427; AAA59576.1; JOINED.  
DR EMBL: M29428; AAA59576.1; JOINED.  
DR EMBL: M29429; AAA59576.1; JOINED.  
DR EMBL: M29430; AAA59576.1; JOINED.  
DR EMBL: M29431; AAA59576.1; JOINED.  
DR EMBL: M29432; AAA59576.1; JOINED.  
DR EMBL: M29433; AAA59576.1; JOINED.  
DR EMBL: M29434; AAA59576.1; JOINED.  
DR EMBL: M29435; AAA59576.1; JOINED.  
DR EMBL: M29436; AAA59576.1; JOINED.  
DR EMBL: M29437; AAA59576.1; JOINED.  
DR EMBL: M29438; AAA59576.1; JOINED.  
DR EMBL: M29439; AAA59576.1; JOINED.  
DR EMBL: M29440; AAA59576.1; JOINED.  
DR EMBL: M29441; AAA59576.1; JOINED.  
DR EMBL: M29442; AAA59576.1; JOINED.  
DR EMBL: M29443; AAA59576.1; JOINED.  
DR EMBL: M29444; AAA59576.1; JOINED.  
DR EMBL: M29445; AAA59576.1; JOINED.  
DR EMBL: M29446; AAA59576.1; JOINED.  
DR EMBL: AF016535; AAB69423.1; -  
DR EMBL: AC002457; AAC82531.1; -  
DR EMBL: M37724; AAB8047.1; -  
DR EMBL: M37725; AAB8048.1; -  
DR EMBL: X58723; CAA41558.1; -  
DR PIR: A25059; DVH01.  
DR PIR: A34914; A34914.  
DR MIM: 171050; -  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001140; ABC\_transportr\_tmemb.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
KW Multigene family.  
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 52 72 POTENTIAL.  
FT TRANSMEM 120 140 POTENTIAL.  
FT TRANSMEM 189 209 POTENTIAL.  
FT TRANSMEM 216 236 POTENTIAL.  
FT TRANSMEM 297 317 POTENTIAL.  
FT TRANSMEM 326 346 POTENTIAL.  
FT DOMAIN 347 710 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 711 731 POTENTIAL.  
FT TRANSMEM 757 777 POTENTIAL.  
FT TRANSMEM 833 853 POTENTIAL.  
FT TRANSMEM 854 874 POTENTIAL.  
FT TRANSMEM 937 957 POTENTIAL.  
FT TRANSMEM 974 994 POTENTIAL.  
FT DOMAIN 1280 1280 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 427 434 ATP (BY SIMILARITY).  
FT NP\_BIND 1070 1077 ATP (BY SIMILARITY).  
FT REPEAT 1 637 ATP (BY SIMILARITY).  
FT REPEAT 638 1280  
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 23 23 S -> R (IN REF. 6).  
FT CONFLICT 185 185 G -> V (IN REF. 1 AND 3).

FT CONFLICT 336 336 MISSING (IN REF. 3).  
FT CONFLICT 412 412 G -> A (IN REF. 3).  
FT CONFLICT 438 438 G -> S (IN REF. 3).  
SQ SEQUENCE 1280 AA; 141462 MW; ABIC27931F43675 CRC64;

Query Match 56.8%; Score 1555.5; DB 1; Length 1280;

Best Local Similarity 55.2%; Pred. No. 1.1e-99; Mismatches 99; Indels 49; Gaps 6;

Matches 306; Conservative 100; Mismatches 99; Indels 49; Gaps 6;

QY 1 MILGILASLVNGACPLPLMLVLGEMSD-----NLISGLVQV--MTYSFF----- 43  
Db 51 MYVGTLLAIIHGAGLPLMLVFGEMTDIPANNGNLEDLSNTRSDINDTGPFNMLEED 110  
QY 44 --RLIIYVIGVALLFETIOTISLITTAQTKRIKROFHSVLAQIDGIFDSCDICE 101  
Db 111 MTRVYVYVSGIAGVVAAYIYVFWCLAAQOIHRIKROFHAIRKQIFGMDVDVGE 170  
QY 102 LNTWMT-DIDKISDGIIDGIALIFQMSSTFISGLAVGLKMKLPLVLTSTPLIMASAA 160  
Db 171 LNTRLTDVSKINEGIGDKIGMFGQSMATFFPGFYGTFRGKRLTLVILAIISVGLSAA 230  
QY 161 ACSRWVLSLSKELSAVSAAGVAEEVLSIRTVIAFRAQEKELORSFLLNTRVAMEYF 220  
Db 231 VVAKILSFTDKELALAYAKAGVAEEVLAIRTVIAFGQKKELE----- 275  
QY 221 PGMILSCVLAFFRYTONLKDAKDFGKRTIASKVSLGAYFFPMNGYGLAFWYGTSLIN 280  
Db 276 -----RYNKNLEAKRIGIKKAITANISGAAPFLIYASALAFWGTTLVLS 323  
QY 281 GPGYTGIVLAVFVSIVHSYCIGAAPHFETFAIRGAAPHFOVIDKPSIDNFSTA 340  
Db 324 GE--YSIGVLYVFSVILGAFSVGQASPSIEFAARAKRAAEIPIIDNKRSIDYSKS 381  
QY 341 GYKPEISIEGVEFKVNSFNPSPRSIKILKGLNLRKISGETVALVGLNCGKSTVQVLQ 400  
Db 382 GHKPKIKNLEFRNVHFYPSRKEVKILKGLNLKQVQSGTVALVGLNCGKSTVQVLQ 441  
QY 401 RLYDDGDFIMDENDIRALNVRHNRDHIGVYSQEVLEGTITISNNIKGRDDVTEEME 460  
Db 442 RLYDDPEGVSVGDODIRTNVFRLEIIGVYSQEVLEPATYIAENIRGRENVYDEDE 501  
QY 461 RAAREANAYDFIMEPNKNTLVGEKQAGMSGOKORIRIARALVNPILILDEATSL 520  
Db 502 KAVKEANAYDFIMKLPKHPKDTLVGERQADLSGQKORIRIARALVNPILILDEATSL 561  
QY 521 DSESKSAVOALEK 534  
Db 562 DRESEAVQVADLK 575  
RESULT 3  
MDR2\_CRIGR STANDARD; PRT; 1276 AA.  
ID MDR2\_CRIGR  
AC P21449;  
DT 01-MAY-1991 (Rel. 18. Created)  
DT 01-NOV-1991 (Rel. 20. Last sequence update)  
DT 01-NOV-1997 (Rel. 35. Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).  
GN PGY2 OR PGP2.  
OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
OX NCBI\_Taxid=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92135896; PubMed=1685679;  
RA Endicott J.A., Sarangi F., Ling V.;  
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein  
gene family.";  
RL DNA Seq. 2:89-101(1991).  
RN [2]

```

RP      SEQUENCE OF 622-1276 FROM N.A...
RA      MEDLINE=88122132; PubMed=2893255;
RA      Endicott J.A., Juranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
RA      Ling V.;
RT      "Simultaneous expression of two P-glycoprotein genes in
RT      drug-sensitive Chinese hamster ovary cells.";
CC      Mol. Cell. Biol. 7:4075-4081(1987).
CC      CC
CC      -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC      DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC      -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC      CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC      CANNOT.
CC      -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC      (ABC TRANSPORTERS). MDR SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.ch).
CC      -----
DR      EMBL: M60041; AAA68884.1; -.
DR      EMBL: M17896; AAA37007.1; -.
DR      PIR: B27126; DVH2C.
DR      InterPro: IPR003593; AAA.
DR      InterPro: IPR003439; ABC_transport.
DR      InterPro: IPR001140; ABC_transport_tmemb.
DR      InterPro: IPR001687; ATP_GTP_A.
DR      Pfam: PF00664; ABC_membrane; 2.
DR      Pfam: PF00005; ABC_tran; 2.
DR      SMART: SMO0382; AAA; 2.
DR      PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW      ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
DR      MultiGene family.
DR      KM
DR      FT      DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
DR      FT      TRANSSEM 52 72 POTENTIAL.
DR      FT      TRANSSEM 119 139 POTENTIAL.
DR      FT      TRANSSEM 188 208 POTENTIAL.
DR      FT      TRANSSEM 215 235 POTENTIAL.
DR      FT      TRANSSEM 296 316 POTENTIAL.
DR      FT      TRANSSEM 325 345 POTENTIAL.
DR      FT      DOMAIN 346 708 CYTOPLASMIC (POTENTIAL).
DR      FT      TRANSSEM 709 729 POTENTIAL.
DR      FT      TRANSSEM 755 775 POTENTIAL.
DR      FT      TRANSSEM 831 851 POTENTIAL.
DR      FT      TRANSSEM 852 872 POTENTIAL.
DR      FT      TRANSSEM 935 955 POTENTIAL.
DR      FT      TRANSSEM 972 992 POTENTIAL.
DR      FT      DOMAIN 993 1276 CYTOPLASMIC (POTENTIAL).
DR      FT      NP_BIND 426 433 ATP (POTENTIAL).
DR      FT      NP_BIND 1068 1075 ATP (POTENTIAL).
DR      FT      REPEAT 1 635
DR      FT      REPEAT 636 1276
DR      FT      SEQUENCE 1276 AA; 141057 MW; 5096B1385628812D CRC64;
QY      Query Match 56.8%; Score 1555; DB 1: Length 1276;
QY      Best Local Similarity 54.7%; Pred. NO.1.2e-99;
QY      Matches 305; Conservative 99; Mismatches 96; Indels 58; Gaps 6;
Db      1 MILGILASLVNGACDPLMPLVGEKSDN-----LISGCLVQ-T 37
Db      51 MWLGILAVLVHNTSLPLMLVPGNNITDSFKAEISWPMNTNOSLNNTEVVISGSLSEDM 110
QY      38 NTYSFRLLYVYGIGVAALLFGYIQISLWITTAARQTKRKIRKQFHSVLAODIGWDSG 97
Db      111 ATYAA-----YTTGIGAGVLIVAYIQVSPWCCLAARQINKRKQFFHAIINQDEIGWDVH 165
QY      98 DEIGLNTNMT-DIDKISGIGDKIALIQNMSTESTIGLAVGKWKMLTVLTLSPFLM 156

```

DB	166	DIGELNTRITDDVSKINDGDKIGMFQSIATFLAFTVIGFISQMKLTLLVLAASPLIG	225
OY	157	ASAAACSRMVISLTSKELSAISKAAVADEVLSIRTYIAPRAQKELORSFLNITRYA	216
DB	226	LSSAAAMVVLTSFTKELQAAKAAVAEEVLAARVIAAGGQKKELE-----	274
OY	217	WFFPWILLSCVLXEVRYTONLKADKDGKIRKTIASKVSLGAVFPFMNGTGLAFWYGS	276
DB	275	-----RINKNLEKKNNGIKKAVYANISIGAVYLLVYASVALAWYGS	318
OY	277	LILNCEPGVITGTVLAVFEFVSIISSYICGAAPVHFETPALRGAAPHFOYIDKPSIDN	336
DB	319	LYLSNE--YVGQVLTVEFSILFTGFSIGHIAPNIEVEFAMNARGAAYETFKIIDNEPSIDS	376
OY	337	PSTAGYKRESIIGVYEFKFNVSFNYSRPSIKIKLNLIRKSGEVALVGLNGSKSTVY	396
DB	377	FSTQHKRDSVNGNLEFNVHFSYPSRSGIKLKGLNKLVGSQGVYLVAGSGCKSTTV	436
OY	397	OLLOKLYPDDGFIVWDENDIRALVRHVRDHIGVVSQEPVLFGTTSNNIRKYGRDVTD	456
DB	437	OLLOKLYPDTESVSIIDGODITRTVRLKRELTIGVVSQEPVLFATTIENIRYGENVTM	496
OY	457	EEEMFAAREANAYDIMEFENKFNLTIVGEKQAGMSGQKORIAIARALVRNPKLILDEA	516
DB	497	DEIEKAVYEAANNAYDIMEFENKFNLTIVGEKQAGMSGQKORIAIARALVRNPKLILDEA	556
OY	517	TSALDESSEKSAVOALEK	534
DB	557	TSALDESEAVVQOALDK	574
RESULT	4		
ID	MDRI_RAT	STANDARD:	PRT: 1277 AA.
AC	P43245:		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).		
GN	ABCB1 OR PGY1 OR MDRI OR MDR1B.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9203081; PubMed=1682220;		
RA	Silverman J.A., Raunio H., Gant T.W., Thoregerson S.S.;		
RT	"Cloning and characterization of a member of the rat multidrug		
RT	resistance (mdr) gene family.";		
RL	Gene 106:229-236(1991).		
CC	-1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED		
CC	DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.		
CC	-1- PM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE		
CC	LINKER DOMAIN.		
CC	-1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE		
CC	RELATED BUT DISTINCT CELLULAR GENES.		
CC	-1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY		
CC	(ABC TRANSPORTERS). MDR SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL: M81855; -; NOT ANNOTATED_CDS.		
DR	InterPro: IPR003593; AAA.		
DR	InterPro: IPR003439; ABC_transport.		
DR	InterPro: IPR001140; ABC_transport_tmem.		

```

DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF000664; ABC_membrane; 2.
DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 71 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT DOMAIN 348 709 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 710 730 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 832 852 POTENTIAL.
FT TRANSMEM 854 874 POTENTIAL.
FT TRANSMEM 937 957 POTENTIAL.
FT TRANSMEM 968 988 POTENTIAL.
FT DOMAIN 989 1277 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (BY SIMILARITY).
FT NP_BIND 1070 1077 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1277
FT CARBOHD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1277 AA; 141386 MW; 8AFDD619D2934C1 CRC64;

Query Match 56.5%; Score 1548; DB 1; Length 1277;
Best Local Similarity 55.3%; Pred. No. 3.6e-99;
Matches 307; Conservative 94; Mismatches 104; Indels 50; Gaps 5;

```

```

OY 1 MILGILASLVNACPLPMLPVLGEMSDNLI-----SGCLVOTNRYSPFL-- 45
DB 49 MALGLAIIHETLPLMLLVGYMDSFTPSRDRPHSDAINTQSIINSHYSDSLG 108
OY 46 -----TLVYVGIGVALLFGYIQISLWITTAARQTKRIKOPHVSVAODIGMFDSCD 100
DB 109 DMAMVAYYYTGGAGVLIYVYGVISLWCLAGROIKRIKQEFHAIMNOIGVFDVNDG 168
OY 101 ELNTRMT-DIDISDGIKIALIQQNMSTFSGIAGVLGKGLVLTLSPLIMASA 159
DB 169 ELNTRITDVSXINDGIGKLMFPOSITTFSGAFTIGTSGKTLVLLAVSPLIGLS 228
OY 160 AACSRNVLISLTSKELSAVSKAGAAVEVLSTRTVAFRAQEKELORSFLNTRYAMEY 219
DB 229 AMMAVILTSFTKKELOAVAKAGAAVEVLAIKRTVIAFGGOKKELE----- 274
OY 220 PQOMLSCVLAHVRYTONLKDADGKIKRTIASKVSGLAVYFPFANGTYGLAEWYGTSL 279
DB 275 -----RYNKNLEAKKRGVQIKRAITANISIGIAYLIVYASVLAIFYGTSLV 321
OY 280 NCEPGYIGTVLAVFVSYSYSCGAAPHFETPAIAGAAPHIEQVLDKPSIDNFE 339
DB 322 SNE--YSIGQVLTVEFSILGTFSIGHLAPNTEAFANAGAAELPKIINDSIDSFST 379
OY 340 AGYKESIEGTEVFNKVSFNPSRPSIKILKGLNIRKSGEVAALVGLSGSKSTVQL 399
DB 380 KGHKDSINGNLEFNKVNYSRSEVKILKGLNKLKVGSGQYIALVGNSGCKSTVQL 439
OY 400 ORLYPDGDFIWDENDIALNVRHRIHIGVVSQEPVLFGTTISNNIKYGRDVTDEEM 459
DB 440 ORLYPIEIEVSIDCODITINRYREITIGVVSQEPVLFATTIAINIRYGRENVMTDEI 499
OY 460 EEAARPAANYDIMEPPNKNLTVGKGAOMSGGOKORAIARARANKIILLDRATSA 519
DB 500 EKAARPAANYDIMEPPNKNLTVGKGAOMSGGOKORAIARARANKIILLDRATSA 559
OY 520 LDESEKSAVOALEK 534
DB 560 LDTESAVVQALDK 574

```

```

RESULT 5
ID MDR3_MOUSE STANDARD; PRT; 1276 AA.
AC P21447;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3) (MDR1A).
GN ABCB4 OR PGY-3 OR MDR3 OR MDR1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205845; PubMed=1969610;
RA Devault A., Gros P.;
RT "Two members of the mouse mdr gene family confer multidrug resistance
RT with overlapping but distinct drug specificities."
RL Mol. Cell. Biol. 10:1652-1663(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287150; PubMed=1972547;
RA Hsu S.I.H., Cohen D., Kirschner L.S., Lochstein L., Hartstein M.,
RA Horwitz S.B.;
RT "Structural analysis of the mouse mdr1a (P-glycoprotein) promoter
RT reveals the basis for differential transcript heterogeneity in
RT multidrug-resistant J774.2 cells."
RL Mol. Cell. Biol. 10:3596-3606(1990).
RN [3]
RP SEQUENCE OF 173-1276 FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=89308614; PubMed=2473069;
RA Hsu S.I.H., Lochstein L., Horwitz S.B.;
RT "Differential overexpression of three mdr gene family members in
RT multidrug-resistant J774.2 mouse cells. Evidence that distinct P-
RT glycoprotein precursors are encoded by unique mdr genes."
RL J. Biol. Chem. 264:12053-12062(1989).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M30697; AAA39517.1; -
CC EMBL: M33581; AAA39514.1; -
CC EMBL: M33580; AAA39518.1; -
CC EMBL: M24417; AAA03243.1; -
CC PIR: A34175; DYMSIA.
CC PIR: A34786; A34786.
CC MGI: 97570; Abcb4.
CC InterPro: IPR003593; AAA.
CC InterPro: IPR003439; ABC_transport.
CC InterPro: IPR001140; ABC_transport_tmam.
CC InterPro: IPR001687; ATP_GTP_A.
CC Pfam: PF00664; ABC_membrane; 2.
CC SMART: SM00382; AAA_2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 2.
CC ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.

```

```

FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 71 POTENTIAL.
FT TRANSMEM 117 136 POTENTIAL.
FT TRANSMEM 188 205 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 327 346 POTENTIAL.
FT DOMAIN 347 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 753 773 POTENTIAL.
FT TRANSMEM 829 848 POTENTIAL.
FT TRANSMEM 853 880 POTENTIAL.
FT TRANSMEM 942 961 POTENTIAL.
FT TRANSMEM 964 984 POTENTIAL.
FT DOMAIN 985 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 423 430 ATP (BY SIMILARITY).
FT NP_BIND 1066 1073 ATP (BY SIMILARITY).
FT REPEAT 1 632
FT REPEAT 633 1276
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 526 527 OL -> HY (IN REF. 2 AND 3).
FT CONFLICT 939 939 F -> S (IN REF. 2 AND 3).
FT CONFLICT 1036 1036 F -> V (IN REF. 2 AND 3).
SQ SEQUENCE 1276 AA; 140754 MW; 75C71F3E1F58481 CRC64;

```

```

Query Match 56.3%; Score 1542; DB 1; Length 1276;
Best Local Similarity 53.8%; Pred. No. 9.2e-99;
Matches 299; Conservative 103; Mismatches 98; Indels 56; Gaps 5;

```

```

QY 1 MILGILASLVACPLPLPLVLEMSDNLISGLVOTN-----T 39
D 50 MLVGLAIHIGVALPLMLLIFGDMTDFASVGNVSKNSTNSSEADKRAMFAKLEEMWT 109
QY 40 YSFRRLTYVYGICVALLIFGIQISLWITPAKOTKRIKROFHVSVALAODIGFDSCDI 99
D 110 YAY-----YTGICGAGVILVAVIQVSEFCLAAAGROIHRIKQFHAHNMEOIGPDDVHD 164
QY 100 GELNTRMT-DIDKISDGDIGDITALLFQNMSTFISGLAVGVKWLTLVTLSPLIMAS 158
D 165 GELNTRLDDVSKINEGIGDKIGMFEQAMTFEGGFIIGTRGWLTLVTLAISPLIGLS 224
QY 159 AAACSRWVISTSKELSLYSKAGAAVEVLSIRTVIAFRAOERKLORSFLLNTRYAMF 218
D 225 AGIYAKILISFTDKELHAYAKAGAAVEVLAIRTVIAFGGCKELE----- 271
QY 219 YFPQWLSCVLXFVRYTONLDAKDEGIKRTIASKVSLGAVYFFNNGTYGLAFWYGTSLI 278
D 272 -----RYNNNLEAKRLGIKAITANISMGAAFLLIYASVALAFWYGTSLV 317
QY 279 LNGERGTYIGTVANFEFVSIHSSYICGAAPHEFTFATARGAAPHIIOVDIKKPSIDNFS 338
D 318 ISKE--VSIGVLYVFEFVLVIGAESVGAQSPHIEAFANAKGAAVEFKIIDKKPSIDFS 375
QY 339 TAGYKPEISIECTVEKESFNVPSPRSIKILKGLNLRIKSGEFTVALVGNSSGKSTVVL 398
D 376 KSGHKPDIOGNLEKRNHFSYPSKKEVOILKGLNLKYSQGTVALVGNSSGKSTVYQL 435
QY 399 LQRLVDPDGLMVDENDIRALNVRHRYDHIGVSOEPVLEFGTTISNNIKYGRDDVDEE 458
D 436 MQRLLDPLDGWMSIDGQIRITINRYLREIRIGVSOEPVLEFATTAENLRGREDVTWDE 495
QY 459 MERARERANVDFIMEFNRKNTLVGEGKAGMSGOKORIAARLVNPKLLIIDETS 518
D 496 TEKAVKEANAADFINKLHGFDTLVGERGAOLSGGOKORIALARLVNPKLLIIDETS 555
QY 519 ALDSEKSAVOALEK 534
D 556 ALDTESEAVVOALDK 571

```

```

RESULT 6
ID MDRL_CRIGR STANDARD; PRT: 1276 AA.
AC P21448;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).
OS ABCB1 OR PGP1 OR PGP1.
NC Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
RT gene family.";
RL DNA Seq. 2:89-101(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91154265; PubMed=1671863;
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced pgp1 transcripts in multidrug-
RT resistant Chinese hamster lung cells.";
RL J. Biol. Chem. 266:4545-4555(1991).
RN [3]
RP SEQUENCE OF 706-1276 FROM N.A.
RX MEDLINE=88122132; PubMed=2893255;
RA Endicott J.A., Juranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
RA Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
RT drug-sensitive Chinese hamster ovary cells.";
RL Mol. Cell. Biol. 7:4075-4081(1987).
CC -I- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
CC EMBL: M60040; AAA68883.1; -
CC EMBL: M59253; AAA37004.1; -
CC EMBL: M17897; AAA37006.1; -
CC PIR: A38696; DVHYIC.
CC InterPro: IPR003593; AAA.
CC InterPro: IPR003439; ABC_transport.
CC InterPro: IPR001140; ABC_transport_tmem.
CC InterPro: IPR001687; ATP_GTP_A.
CC Pfam: PF00664; ABC_membrane; 2.
CC Pfam: PF00005; ABC_tran; 2.
CC SMART: SM00382; AAA; 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 2
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 50 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.

```

```

FT TRANSMEM 323 343 POTENTIAL.
FT DOMAIN 344 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 754 774 POTENTIAL.
FT TRANSMEM 830 850 POTENTIAL.
FT TRANSMEM 851 871 POTENTIAL.
FT TRANSMEM 934 954 POTENTIAL.
FT TRANSMEM 971 991 POTENTIAL.
FT DOMAIN 992 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 424 431 ATP (POTENTIAL).
FT NP_BIND 1067 1074 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1276
FT CONFLICT 338 339
FT SEQUENCE 1276 AA; 140925 MW; 44F3F92A186BADFF CRC64;

Query Match 56.3%; Score 1541.5; DB 1; Length 1276;
Best Local Similarity 54.0%; Pred. No. 1e-98;
Matches 301; Conservative 100; Mismatches 99; Indels 57; Gaps 5;

OY 1 MILGILASLVAGACLPMLPLVLGEMSDNLISGLVOTN----- 38
DB 50 MLVGTLLAIHIGVALPLMLLVFGDMTDFASVGNIPNTATNATQVNASDIKLEEMT 109
OY 39 TYSFRLLYVYGVVALIFGYTOISIMITTAQOTRIKQFPHSVLAODIGMFDSCD 98
DB 110 TYAV---YVTGAGVLAIVYIOVSFCLAGROIKIKROKPHALMOEIGMFDVMD 164
OY 99 IGLNTMT-DIDKISDIDGKIALLFONMSTFISGLAVGLKGMKLTLTSTPLIVA 157
DB 165 VGLNTRITLDVSKINEIGDKIGFPMATFFGFLIIGTRGMKLTLTIAISPVLG 224
OY 158 SAACSRNVISITSKELSAVSKAGAVAEVYSSITVTIAFRAQEKELORSPLNITRAM 217
DB 225 SAGIWAAILSSFTDKELQAVAKAGAVAEVLAIRTVIAFGQKKELE----- 272
OY 218 FYFPMWLSCVLXFRYRQNKDKADQGIKRTIASKVSGANVFFPMNGYGLAFWYGTSL 277
DB 273 -----RYNNLEEARLRIKAITANISGAFLIIVASYALAFWYGTSL 317
OY 278 ILNGEPYTGTVLAFVFSVSHSYCIGAANVPHETFAIARGAAPHFOVIDKPSIDNF 337
DB 318 VISKE--YSIGOVLFVFAVLGAFSIGOASNIIEAFANARAALIEIFIIDNKSIDSF 375
OY 338 STAGYKPSISIGYVEFKVNSFNPSPRSIKILKGLNLRIKSGEVALVGLNGSKSTVVO 397
DB 376 SKNGYKPNIKGNLEFKNIHESYPSRKDVQILKGLNLKVGQFVALVGNSCGKSTVVO 435
OY 398 LLQRIYDDDFIMVDENDIRALNRIHRDHIGVYSOEPLYEGTTISNNIKYGRDVTDE 457
DB 436 LLQRIYDPEGVASTIDGQDIFTINRYRIEIIIGVYSOEPLYEATTAIEIRIGREVYMD 495
OY 458 EMERARARANAVDFIEMEPNKFNLTVEGGAOMSGOKORIRIARALVNPFIILLIDEAT 517
DB 496 ELEKAVKANAYDFIMKLPKFTDLVGERGALSGGOKORIRIARALVNPFIILLIDEAT 555
OY 518 SALDESSESAVOALEK 534
DB 556 SALDESSESAVOALDK 572

RESULT 7
MDR3_HUMAN STANDARD: PRT: 1279 AA.
ID_MDR3_HUMAN
AC P21439;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3).
GN ABCB4 OR PGY3 OR MDR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; PubMed=2906314;
RA van der Biele A.M., Koolman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL Gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=8811519; PubMed=2892668;
RA van der Biele A.M., Baas F., ten Houte de Lange T., Koolman P.M.,
RT "van der Biele-Koerts T., Borst P."
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL Embo J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; PubMed=2002063;
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
CC -I- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT
CC CAPABLE OF CONFERRING DRUG RESISTANCE. MEDIATES THE TRANSLLOCATION
CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
CC HEPATOCYTE.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- DISEASE: DEFECTS IN PGY3 ARE A CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS TYPE IIT (PEIC), A FORM OF AUTOSOMAL
CC RECESSIVE LIVER DISORDERS, CHARACTERIZED BY EARLY ONSET OF
CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
CC ADULTHOOD.
CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23234; AAA36207.1; -.
DR EMBL: X06181; CAA29547.1; -.
DR PIR: JS0051; DVHJ3.
DR PIR: AA2213; A42213.
DR HSP: P13569; 1NBD.
DR MIM: 171060; -.
DR MIM: 602347; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABC_transporter_tmern.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT TRANSMEM 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 354 POTENTIAL.
FT TRANSMEM 355 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.

```







```

Db 231 AAVMAKILSTFSDKELAAKAGAAVEELGAIIRVIANGGONKELE----- 277
Qy 219 YFPQWLSCVLAFFVRYTONLAKADQGIKRTTASKVSICGAVFFNMGTGGLAFWGTSLI 278
Db 278 -----RYOKHLENNAKKIGIKKAIKSAISMWIAFLLIYASALAFWGSTLY 323
Qy 279 LINGEPYITIGTVLAVFEFVSHSSYCGAANVPHFEPFAIRGAFAHIFQYIDKKPSIDNS 338
Db 324 ISKE--YITGNAMTVFFSLIGAFSGOAPCIDAFANRGAAYVIEDIIDNNPKIDSFS 381
Qy 339 TAGYRPESIEGTEFEKKNVSFNPSRPSIKILKGLNIRIKSGTEVALVNGSGKSTVYL 398
Db 382 EGHKRPDSIKGKLDFSDVAFSYPSTRANIKILKGLNIAKVSOGTVALVNGSGCKTTLDL 441
Qy 399 LQRLYDPDDGFLMVDENDIRALNVHRHYDHIGVVSQEPVLFSTTISNNIKYGRDVTDEE 458
Db 442 LQRLYDPDEGTISIDGDIRNFNVRCLREIIGVVSQEPVLFSTTIAENIRYGRGNTVME 501
Qy 459 MERAAREANAYDFIMEFPKFFNTLVGEKAGMSGGOKRIATARALVRRPKILLDEANS 518
Db 502 IKKAVKEANAYDFIMKLPPQKFDTLVGERGAQLSGGOKRIATARALVRRPKILLDEANS 561
Qy 519 ALDSEKSAVQALES 534
Db 562 ALDTESEAEVQAALDK 577

RESULT 9
MDR2_MOUSE STANDARD: PRT: 1276 AA.
ID P21440:
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).
GN PGY2 OR PGY-2 OR MDR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88302195; PubMed=3405218;
RA Gros P., Raymond M., Bell J., Housman D.;
RT "Cloning and characterization of a second member of the mouse mdr
RT gene family.";
RL Mol. Cell. Biol. 8:2770-2778(1988).
RN [2]
RP SEQUENCE OF 1-23 FROM N.A.
RC STRAIN=BAIB/C;
RA Kirschner L.S., Horwitz S.B.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS
CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J03398; AAA39516.1;
CC DR EMBL: M74151; AAA39515.1;
CC PIR: A30409; DVMS2.

```

```

DR HSSP: P13569; 1NBD.
DR MGD: MGI:97569; Pgy2.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABC_transporter-tmem.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00605; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Domain family.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 75 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 707 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 708 728 POTENTIAL.
FT TRANSMEM 753 773 POTENTIAL.
FT TRANSMEM 829 849 POTENTIAL.
FT TRANSMEM 850 870 POTENTIAL.
FT TRANSMEM 933 953 POTENTIAL.
FT TRANSMEM 970 990 POTENTIAL.
FT DOMAIN 991 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (BY SIMILARITY).
FT NP_BIND 1066 1073 ATP (BY SIMILARITY).
FT REPEAT 1 635
FT REPEAT 636 1276
SQ SEQUENCE 1276 AA; 140332 MW; A6C38DCD2C118EEF CRC64;

```

```

Query Match 52.6%; Score 1439.5; DB 1; Length 1276;
Best Local Similarity 51.3%; Pred. 1.1e-91;
Matches 285; Conservative 103; Mismatches 111; Indels 57; Gaps 6;

Qy 1 MIIIGTILSVNGACPLMPLVIGEMSDNLSGCLVQNTYSF-----FRLLT----- 47
Db 54 MFLGTLMAIAHGSGLPLMNIIVFGEKMDKFDV-----NTGNFSLPVNFSLISMNIPGRILE 107
Qy 48 -----YVVGIGVALAIFGYIQTSLMTITTAARQTKRIKQFHSVLNODIWMPSDOI 99
Db 108 EEMTRYAYVYSGGGLVAVATVQVFWTLAAGQIKTKRQKFAHLRQEMKMPDIKGT 167
Qy 100 GELNTRMT-DIDKISDIDKIALLFQNNSTFSGIAGVLGKGMKLLVTLSTSPIMAS 158
Db 168 TELNTRFLTDVSKISISGIDGKVMFQATATPAGTIVGIRGMKLLVIMAISSLTIGLS 227
Qy 159 AAACSRMTVLSLTSKELSAVSKAGAAVEVLSSIRVIAPRAQEKELORSFLNTRYAMP 218
Db 228 TAVMAKILSTFSDKELAAKAGAAVEEAPGAIIRVIANGGONKELE----- 274
Qy 219 YFPQWLSCVLAFFVRYTONLAKADQGIKRTTASKVSICGAVFFNMGTGGLAFWGTSLI 278
Db 275 -----RYOKHLENNAKKIGIKKAIKSAISMWIAFLLIYASALAFWGSTLY 320
Qy 279 LINGEPYITIGTVLAVFEFVSHSSYCGAANVPHFEPFAIRGAFAHIFQYIDKKPSIDNS 338
Db 321 ISKE--YITGNAMTVFFSLIGAFSGOAPCIDAFANRGAAYVIEDIIDNNPKIDSFS 378
Qy 339 TAGYRPESIEGTEFEKKNVSFNPSRPSIKILKGLNIRIKSGTEVALVNGSGKSTVYL 398
Db 379 EGHKRPDSIKGKLDFSDVAFSYPSTRANIKILKGLNIAKVSOGTVALVNGSGCKTTLDL 438
Qy 399 LQRLYDPDDGFLMVDENDIRALNVHRHYDHIGVVSQEPVLFSTTISNNIKYGRDVTDEE 458
Db 439 LQRLYDPDEGTISIDGDIRNFNVRCLREIIGVVSQEPVLFSTTIAENIRYGRGNTVME 498
Qy 459 MERAAREANAYDFIMEFPKFFNTLVGEKAGMSGGOKRIATARALVRRPKILLDEANS 518
Db 499 IKKAVKEANAYDFIMKLPPQKFDTLVGERGAQLSGGOKRIATARALVRRPKILLDEANS 558

```



CC - FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
 CC INTO THE CANALICULUS OF HEPATOCYTES.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
 CC SITU (BY SIMILARITY).  
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
 CC ATP BINDING CASSETTE (ABC) DOMAIN.  
 CC -1- DISEASE: DEFECTS IN ABCB11 ARE THE CAUSE OF PROGRESSIVE FAMILIAL  
 CC INTRAHEPATIC CHOLESTASIS 2 (PEIC2), AN INHERITED LIVER DISEASE OF  
 CC CHILDHOOD. PEIC2 IS CHARACTERIZED BY CHOLESTASIS AND NORMAL SERUM  
 CC GAMMA-GTAMYLTRANSFERASE ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF091582; AAC77455.1; -  
 DR EMBL: AF136523; AAD28285.1; -  
 DR MIM: 603201; -  
 DR MIM: 601847; -  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001140; ABC\_transport\_tmem.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 DR ATP-binding; Transmembrane; Transport; Disease mutation.  
 KW DOMAIN 1 62  
 FT TRANSMEM 63 83  
 FT DOMAIN 84 147  
 FT TRANSMEM 148 168  
 FT DOMAIN 169 215  
 FT TRANSMEM 216 236  
 FT DOMAIN 237 240  
 FT TRANSMEM 241 261  
 FT DOMAIN 262 319  
 FT TRANSMEM 320 340  
 FT DOMAIN 341 353  
 FT TRANSMEM 354 374  
 FT DOMAIN 375 755  
 FT TRANSMEM 756 776  
 FT DOMAIN 777 794  
 FT TRANSMEM 795 815  
 FT DOMAIN 816 869  
 FT TRANSMEM 870 890  
 FT TRANSMEM 891 911  
 FT DOMAIN 912 979  
 FT TRANSMEM 980 1000  
 FT DOMAIN 1001 1011  
 FT TRANSMEM 1012 1032  
 FT DOMAIN 1033 1321  
 FT NP\_BIND 455 462  
 FT NP\_BIND 1113 1120  
 FT CARBOHYD 109 109  
 FT CARBOHYD 116 116  
 FT CARBOHYD 122 122  
 FT CARBOHYD 125 125  
 FT VARIANT 297 297  
 FT CONFLICT 339 339  
 FT CONFLICT 444 444  
 FT SEQUENCE 1321 AA; 146392 MW; D444CACCA48DEA71 CRC64;

Query Match 45.6%; Score 1250; DB 1; Length 1321;  
 Best Local Similarity 46.0%; Pred. No. 1,4e-78;  
 Matches 262; Conservative 93; Mismatches 151; Indels 64; Gaps 5;  
 1 MIIIGLASLVNACGCLPLMPLVIGMSDNLI-----SCLVQFNT----- 39  
 DB 62 MYVGSICATLHIAQPGVLLIFGTMDFIDVDVLEQIQKACVNNITWTSNLQ 121  
 QY 40 -----YSEFLTYVYGIVALLFGYIQISLIITTAQRTKRIKROFHS 85  
 DB 122 NMTNGRCGLNIESEMIKFASTYAGIAVALITGICFWIAAARQIKRKRYFR 181  
 QY 86 VLAQDIGWDSODIGELNRMF-DIDKISDGIDKALLFQNMSTFISGLAIVKGM 144  
 DB 182 IIRMEIGWFDSCVSELNFRFSDINKINDAIDADQALFIQRTSTICGFLGFE 241  
 QY 145 TLVLTSTPLMASAACSRAVYISLTKELSAVSKAGAAVEVLSIRTVIAFRA 204  
 DB 242 TVIIVSVSLIGATIGATIGLSKFTDYELKAYARAGVADEVISSKRTVA 301  
 QY 205 QRSFLNITRYAMFYFPQWLLSCVLFVRYTQNLKDADFGIKRTASKVSGA 264  
 DB 302 E-----RYEKNLYFAQKMGIRKGIWVGFTGFWCLIF 334  
 QY 265 GTYGLAFWGTSLILNGEGYITIGYLAFFSVYHSYTCGAAPFFETFAIRGA 324  
 DB 335 LCYAAFWYSGTLVDV-BEETPGTLVQIFLSYVAGALMGNAPOCLEAFAGRA 393  
 QY 325 FOVIDKKPSIDNFTAGYRESIEGVEKKNVSPNRSRPIKILKGLRISG 384  
 DB 394 FETIDRKPIIDCMSEDEGKILDKIKETIEFHNVTFHPSPEVKILNDLMV 453  
 QY 385 VGLNCSGKSTVQLQRLVDPDGEIMVDENDIRALNVHRDHTIGVSOEPL 444  
 DB 454 VQPSGAGKSTALQIGRFVDPCEGMYVDGHDIRSINQIMLRQIQIVDEP 513  
 QY 445 NNIKIGRDVYTDENERRAREANADFIEMFPKFTVIGEGAGMSGOKRIAT 504  
 DB 514 ENIRGRGREDATMEDIVQAKKANAYNFIMDLPOQFDTLVEGGGMSGOK 573  
 QY 505 VKNPKILIDENATSAIDSEKSAVOALEK 534  
 DB 574 INPKILLDMATSAIDNESEAMQEVLSK 603  
 RESULT 12  
 AB11\_RAT STANDARD; PRT: 1321 AA.  
 AC 070127;  
 DT 20-AUG-2001 (Rel. 40; Created)  
 DT 20-AUG-2001 (Rel. 40; Last sequence update)  
 DT 20-AUG-2001 (Rel. 40; Last annotation update)  
 DE BLUE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)  
 DE (SISTER OF P-GLYCOPROTEIN).  
 GN ABCB11 OR BSEP OR SPGP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
 RX MEDLINE=98212048; PubMed=9545351;  
 RA Gerloff T., Scieger B., Hagenbuch B., Madon J., Landmann L., Roth J.,  
 RA Hofmann A.F., Meier P.J.;  
 RT "The sister of P-glycoprotein represents the canalicular bile salt  
 RT export pump of mammalian liver";  
 RL J. Biol. Chem. 273:10046-10050(1998).  
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
 CC INTO THE CANALICULUS OF HEPATOCYTES.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
CC SITU.  
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
CC -1- ATP BINDING CASSETTE (ABC) DOMAIN.  
CC SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch)).  
CC -----  
CC EMBL: 069487; AAC40084.1; -.  
CC InterPro: IPR003593; AAA.  
CC InterPro: IPR003439; ABC\_transport.  
CC InterPro: IPR001140; ABC\_transport\_tmam.  
CC InterPro: IPR001687; ATP\_GTP\_A.  
CC Pfam: PF00664; ABC\_membrane; 2.  
CC SMART: SM00382; AAA; 2.  
CC PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
CC ATP-binding; Transmembrane; Transport.  
CC KM  
FT TRANSMEM 1 62 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 63 83 POTENTIAL.  
FT TRANSMEM 84 147 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 148 168 POTENTIAL.  
FT TRANSMEM 169 215 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 216 236 POTENTIAL.  
FT TRANSMEM 237 240 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 241 261 POTENTIAL.  
FT TRANSMEM 262 319 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 320 340 POTENTIAL.  
FT TRANSMEM 341 353 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 354 374 POTENTIAL.  
FT TRANSMEM 375 755 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 756 776 POTENTIAL.  
FT TRANSMEM 777 794 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 795 815 POTENTIAL.  
FT TRANSMEM 816 869 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 870 890 POTENTIAL.  
FT TRANSMEM 891 911 POTENTIAL.  
FT TRANSMEM 912 979 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 980 1000 POTENTIAL.  
FT TRANSMEM 1001 1011 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1012 1032 POTENTIAL.  
FT TRANSMEM 1033 1321 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 1321 455 ATP (POTENTIAL).  
FT TRANSMEM 455 462 ATP (POTENTIAL).  
FT TRANSMEM 462 1120 ATP (POTENTIAL).  
FT TRANSMEM 1120 1133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TRANSMEM 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TRANSMEM 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TRANSMEM 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TRANSMEM 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TRANSMEM 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT TRANSMEM 1321 146257 MW; 5443F4E7B9F1F6 CRC64;  
SQ SEQUENCE

Query Match 44.5%; Score 1220; DB 1; Length 1321;  
Best Local Similarity 44.7%; Pred. No. 1,7e-76;  
Matches 255; Conservative 98; Mismatches 153; Indels 64; Gaps 5;  
OY 1 MILDGLASVAGACLPMPVLEGMSDNL-----SCGLVOTNTY----- 40  
D6 62 MLMGGVALLHGMQPGILIFIGIMTDIFIRQLEIPGRACVNNVTIWNSSFHQ 121  
OY 41 -----SFRRLTYVYGIGVALLIFGYIOISLMTITAAQTRKIRQFHS 85  
D6 122 NMTGCTVCGLYDISEMTEKFSGITAGVGMTYLLIGYFQIRLWITGAIQIRMRKIYRR 181

OY 86 VLAQDQWEDSDQIGELNTRMT-DIDKISDGDKIALLFQNNSTFSGIAGLVAGMKL 144  
D6 182 IMHREGEWEDCTSVSGELNSRFDIDKINDATADQLAHFLQNRSTAMCGILGFGYMKL 241  
OY 145 TLVTLSTPLIMASAAACSRMVISLTSKELSAVSKAGAAVEVLSIRVIAFRAQEKEL 204  
D6 242 TLVTLVAVPLIGAVIGAVIGISIAKFELELKAVAKGSIADDEVLSIRVIAFRAQEKEL 301  
OY 205 QNSFLNITRYAMFPPQWLLSCVLFVRYTQNLKAKDQKIRITASKVSLGAVFFEMN 264  
D6 302 E-----RYEKNLVQAQRMGIRKGMVMEFTYMCCLIF 334  
OY 265 GYGLAFWYGTSLILNGEPYTGIVLAFPSYHSSCGAAPPHEPAIRGAFAFI 324  
D6 335 FCYALFWGSGTVLVD-EERYPTGLVQIFLCVILAMNIGHSSCLFISTGCSATNI 393  
OY 325 FOYIDKPSIDNFSVAGKPESTEGVEFNVSFNPSPRSIKLGLNLRISGETVAL 384  
D6 394 FQIDRQPVIDCMGSDCYKIDRIKGELEFNHVFHPSRDVYKIDLSMVIKRGFTAL 453  
OY 385 VGLNGSGKSTVOLLQRLYPDDGEIWNDENDIRALNVRHVRDHIGVSGEPVLFGTIS 444  
D6 454 VGSAGACKSTALQILQRFYPCGEMVTLDDIRSLNIRMLRQIGIVEQEPVLFSTIA 513  
OY 445 NNIRKYGDDVTDMEERAAAREANAYDFIMEPNKFNLTVEKGAOMSGGOKRIARAL 504  
D6 514 ENIRFGEADTMDIYVAAADANAYNFIMLPQGFDTLVGBGGOMSGGOKORAIARAL 573  
OY 505 VRNPKIILDEATSAIDSEKSAVQALEK 534  
D6 574 IRNPKIILDMATSAIDNESAPVQEALEK 603

RESULT 13  
AB11\_MOUSE STANDARD; PRT; 1321 AA.  
ID AB11\_MOUSE  
AC 090Y30; 090Z8; 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)  
DE (SISTER OF P-GLYCOPROTEIN).  
GN ABCB11 OR BSEP OR SPCP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=20076398; PubMed=10607905;  
RA Green R.M., Hoda F., Ward K.L.;  
RT "Molecular cloning and characterization of the murine bile salt export  
pump";  
RL Gene 241:117-123(2000).  
RN [2]  
RP SEQUENCE OF 463-635 FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Liver;  
RA Salkar R., Suchy F.J., Ananthanarayanan M.;  
RT "Molecular cloning of mouse liver bile salt export pump (bsep).";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
CC INTO THE CANALICULUS OF HEPATOCYTES.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
CC SITU.  
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
CC ATP BINDING CASSETTE (ABC) DOMAIN.  
CC SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF133903; AAF14372.1; -
DR EMBL; AF186585; AAD56419.1; -
DR MGD; MGI:1351619; Abcb11.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABC_transportr_tm.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam; PF00064; ABC_membrane; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transmembrane; Transport.
FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 63 83 POTENTIAL.
FT DOMAIN 84 147 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 148 168 POTENTIAL.
FT DOMAIN 169 215 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 216 236 POTENTIAL.
FT DOMAIN 237 240 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 241 261 POTENTIAL.
FT DOMAIN 262 319 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 320 340 POTENTIAL.
FT DOMAIN 341 353 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 354 374 POTENTIAL.
FT DOMAIN 375 755 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 756 776 POTENTIAL.
FT DOMAIN 777 794 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 795 815 POTENTIAL.
FT DOMAIN 816 869 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 870 890 POTENTIAL.
FT DOMAIN 891 911 POTENTIAL.
FT TRANSSEM 912 979 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 980 1000 POTENTIAL.
FT TRANSSEM 1001 1011 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1012 1032 POTENTIAL.
FT TRANSSEM 1033 1321 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1321 1321 ATP (POTENTIAL).
FT NP_BIND 455 462 ATP (POTENTIAL).
FT NP_BIND 1113 1120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 481 481 L -> P (IN REF. 2).
FT CONFLICT 633 633 T -> V (IN REF. 2).
SQ SEQUENCE 1321 AA; 146675 MW; 15B5EBF175D32967 CRC64;
Query Match 44.0%; Score 1204; DB 1; Length 1321;
Best Local Similarity 44.8%; Pred. No. 2; le-75;
Matches 256; Conservative 97; Mismatches 152; Indels 66; Gaps 7;
OY 1 MILGILASLVGACLPMPPLVLGEMSD----- 27
DB 62 MEMGVVCLLHGMAOPGMIYFGILDFVEYDIEROELSPGKYCMNNTIWMNSFNO 121
OY 28 NLISG--C-LVQTMNTSFFRLTVYVGIGVALLFGYIQLISLMTTARQTKRIKQFEH 84
DB 122 NMTNTSGDIVDINS-EVIRKSGIYAGVAVLLIGYQIRLMVITGARQIRKMKKFFER 180
OY 85 SYLADIDWPSDIDGELNTRMT-DIKDISGIGKIALLFQNMSTFSIGLAVLGKWK 143
DB 181 RIMREIIGFDCSTSGEINRFSDDINKIDEIADOMALFLQRLSTALSGLLGFGYRGWK 240
OY 144 LTVLTLSPILMASAACSRRVLSLTSKELSAVSKAGAVAEVLSIRTVYAFRAQERE 203

```

```

DB 241 LTVLTLAVSPILGIAAVGLSTVAKFTELELKAAVYKAGSIADDEVLSIRTVYAFGCKENK 300
OY 204 LQSRSLNITRAMPYFPFOWMLSCVLXVRYTONKDKAKDEIKRTISKVSLGAVYFEM 263
DB 301 VE-----REKNIIMFQRMGIMWGMPGFTGYMGLI 333
OY 264 NGTYGLAFWYGNLSLNGEPGYTIGTVLAFVPSVHSSVICAAPVHFEPAIARGAAPH 323
DB 334 FFCYALAFWYGSRLVLD-EGEYTPGTLIQLFLCYLITAAAMNIGMASCLEIFSTGSAASS 392
OY 334 IFVOIDKKPSIDNFSAGYKPESEIGTVEFKNVSNYSPRSIKILKILNLRKSGEYVA 383
DB 393 IFQITDRQPVWDMGSGDGYKIDRIKGEIEFHVHTFHSRPEVKILNLSMVIKRGETTA 452
OY 384 LVGLNGSKSTVOLLQRLYDDPGFIWVDENDIALNVRHRYDHIGVSGEPVLEPTTI 443
DB 453 FVGSSGAKKSTALQLIQRFYDCEGMVTLGDHISLNRMLRDOIGIVEQRPVLFSTTI 512
OY 444 SNNIYGRDDVYDEEMERARANAAYDFIMEFPNKNFTLYGKGAOMSGGOKORAIARA 503
DB 513 AENIRLGRBEATVEMEDIVQAKADANVNFIMALPQQGFDTLVGGGOMSGGOKORVALARA 572
OY 504 LVNRNPKIILDEATSAIDSEKSAVOALEK 534
DB 573 LIRKPKIILLDMATISALDNESAKYOGALNK 603
RESULT 14
AB11_RABIT STANDARD; PRT; 1321 AA.
ID AB11_RABIT
AC 09NOV03:
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)
DE (SISTER OF P-GLYCOPROTEIN).
GN ABCB11 OR BSEP OR SPGP
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;
RA Balasubramanian N.V., Suchy F.J., Ananthanarayanan M.;
RT "Molecular cloning and characterization of rabbit liver bile salt
RT export pump (Bsep/spgp).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
CC INTO THE CANALICULUS OF HEPATOCYTES.
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC - TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
CC SITU.
CC - DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC ATP BINDING CASSETTE (ABC) DOMAIN.
CC - SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF249879; AAF65552.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.

```

```

DR InterPro: IPR001140; ABC_transp_rtmem.
DR InterPro: IPR001687; ATP_GRP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran_2.
DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transmembrane; Transport.
KW DOMAIN 1 62
FT TRANSMEM 63 83
FT TRANSMEM 84 147
FT TRANSMEM 148 168
FT TRANSMEM 169 215
FT TRANSMEM 216 236
FT TRANSMEM 237 240
FT TRANSMEM 241 261
FT TRANSMEM 262 319
FT TRANSMEM 320 340
FT TRANSMEM 341 353
FT TRANSMEM 354 374
FT TRANSMEM 375 755
FT TRANSMEM 756 776
FT TRANSMEM 777 794
FT TRANSMEM 795 815
FT TRANSMEM 816 869
FT TRANSMEM 870 890
FT TRANSMEM 891 911
FT TRANSMEM 912 979
FT TRANSMEM 980 1000
FT TRANSMEM 1001 1011
FT TRANSMEM 1012 1032
FT TRANSMEM 1033 1321
FT TRANSMEM 1455 462
FT TRANSMEM 1109 1120
FT TRANSMEM 1109 116
FT TRANSMEM 116 122
FT TRANSMEM 122 125
FT TRANSMEM 125 125
SQ SEQUENCE 1321 AA: 146376 MW: 457539 FCD6D17A2A CRC64;

Query Match 42.18; Score 1153; DB 1; Length 1321;
Best Local Similarity 43.28; Pred. No. 7.1e-72;
Matches 246; Conservative 93; Mismatches 167; Indels 64; Gaps 5;
1 MIIGIILASLVNAGCLPMPVVLGEMSDNLI-----SGCLVQVNTY----- 40
62 MCGSLCACIHGTAPOGVLLIFGTMVDVFDYDELOELKIPGACVNTTIVWINSLSNQ 121
41 -----SFFRLTYVYGIVGVAALIFGYIQISLMTTAAARQTKRIKQEFHS 85
122 NTNGTRCGLLDIESEMIRFAGYVAGIGIAVLGTGYIQICFWGIAAHQIQMKRSEYERK 181
86 VLAODITWFPSCDIEGLNTRKT-DIDKISDQIGKIALIQNMSTFSGLAVALGVKGMKL 144
182 IMRWGIGWVONSVGKNTFSDVFNKINSSADQIAIFDGMSPFEGFLVGSQWKKL 241
145 TIVTISPLIMASAACSRNVISLTSKELSAVSKAGVAEVSIRTVAFPAOEREL 204
242 TIVTISPLIGALIGLSVSKFTDELKATYAKGASVADDEVISSMRTVAAFGEKKEV 301
205 QRSFLNITRYAMFPQWLISCVLSEFVRYQNLKDAKDFGKRTIASKVSIAVYEFMN 264
302 E-----RYEKNLVFAQRMIRKGIWVGFTGIMWCILF 334
265 GYGLAFWYGTSLILNPGYITIVLAVFVSIVHSYCYICAAVPHETFAIANGAFHI 324
335 FCYALAWYGSKIVLE-EGEYSFGALVOIFLSVITIGALINGNASPCLEAFAGRAASSI 393
325 FOVDKRPISDNFSTAGYKPSISIGTVEKFNVSFNPSPRSIKILKGINLIRKGETVAL 384
394 FETIDKRPIDCMSEDDGKTERIGELFHNVTYHPSRPVKILLNNLSIMKIKGEMTAL 453
385 VGLNGSGKSTFVVOILQRLYPDDEGFIWDENDIRALNRYRDIHGIVVSOEPVLFGTTIS 444

```

```

DB 454 VGPSGAGKSTALQILHREYVPTGEMTVESHDIRSSHIOMLRNQIGIVEDEPVLFFHTTA 513
QY 445 NNIRKYGDDVTDDEMEAREANAYDFIMEFPNKFTLVGEKGAQSGGQRIATARAL 504
DB 514 EKIRYGHEDVTMEDLIQAKEANAYNFIMDLPOOPFDLVGEGGQSGGQKQVATARAL 573
QY 505 VRNPKIILIDEATSAIDSESKSAVQAALER 534
DB 574 IRNPKIILDMATSAIDSESAVQEAALSK 603

RESULT 15
ID MDR4_DROME STANDARD; PRT; 1302 AA.
AC 000449;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN HOMOLOG 49 (P-GLYCOPROTEIN 49).
GN MDR49.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Head;
RX MEDLINE=91304385; PubMed=2072901;
RA Wu C.-T., Budding M., Griffin M.S., Croop J.M.;
RT Isolation and characterization of Drosophila multidrug resistance
RT gene homologs.*;
RL Mol. Cell. Biol. 11:3940-3948(1991).
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC - SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59076; AAA28679.1; -.
DR PIR: A41249; A41249.
DR FlyBase: FBgn0004512; Mdr49.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR001687; ATP_GRP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran_2.
DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 46
FT TRANSMEM 47 67
FT TRANSMEM 117 145
FT TRANSMEM 193 213
FT TRANSMEM 221 240
FT TRANSMEM 301 322
FT TRANSMEM 340 360
FT TRANSMEM 361 734
FT TRANSMEM 735 756
FT TRANSMEM 780 801
FT TRANSMEM 856 876
FT TRANSMEM 878 897
FT TRANSMEM 960 980
FT TRANSMEM 997 1017

```

FT DOMAIN 1018 1302 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 437 444 ATP (POTENTIAL).  
FT NP\_BIND 1094 1101 ATP (POTENTIAL).  
FT REPEAT 1 670  
FT REPEAT 671 1302  
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 697 697 S->N.  
FT VARIANT 712 712 D->N.  
FT VARIANT 952 952 V->I.  
SO SEQUENCE 1302 AA: 142724 MW: 4720C169455C28CD CRC64;

Query Match 36.6%; Score 1002; DB 1; Length 1302;  
Best Local Similarity 39.2%; Pred. No. 1.9e-61;  
Matches 222; Conservative 95; Mismatches 189; Indels 60; Gaps 6;

QY 1 MILGILASLVNGACILPLMLVLGKSDNLISGLVQNTYSEFRLTYVG----- 51  
DB 46 LVSLVLTATASAFIPYPMIYGFTSLVDRTVGVTSSPAFALPMFGGQQLTNASKE 105  
QY 52 -----IGVAALIFG---YIOISLMTITAR-----QTKRIKQFFHSVLAODIG 92  
DB 106 ENNOAIIDDAFAFGISLVGSVAMFLITLAIIDLALNOLIDRIKLFLEAMLRODIA 165  
QY 93 WFDSCDIGELNTRMT-DIDKISDIGDKIALLFQNMSTFSIGLAVGLKGMKLTVTLTST 151  
DB 166 WYDTSSSGSNFASKMTEDDKLEKIGEKIVIVFLIMTFVIGIVSAFYGMKLTIVLSC 225  
QY 152 SPLIMASAAACSRMVISLTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELORSFLIN 211  
DB 226 VPFILATSVARLQSGIAEKELKSYSDAANVEEVFGIRTVFAFSGQEKKE----- 279  
QY 212 ITRYAMFFPQWMLSCVLAFTYTONLKDADFGIKRTIASKVSLSGAVYFFPMNGTYGLAF 271  
DB 280 -----RFGKLLIPAEHTGRKKGLYSGMGNALSMILIIYLCMALAI 318  
QY 272 WYGTSLILN---GEPYTGIVLAVFSPVTHSSYICGAAYPHETFAIARGAFHIFOV 327  
DB 319 WYGTLLILDERDLPDRTVTPAVLYVLEFAVIMGALNGFASPHVEAIAVATAAGQTLFNI 378  
QY 328 IDKPSIDNFTAGYKPSIEGTVEFKNVSPNPSRPSIKILKGLNLRIKSGETVALVGL 387  
DB 379 IDRPQVDPMDEKGNRPENTAGHIREGIRFRYPARPVEILKGLTVDYLPQQTVAFWGA 438  
QY 388 NGSGKSTVVOILLQRLYPDDEGFIWDENDIRALNVRHVRDHIGVVSQEPVLEFGTTISNNI 447  
DB 439 SGCGKSTLIQIMOREFYDEAGSVKLDGRDLFTLANGMWLSQIGVYGQEPVLEFATTIGENI 498  
QY 448 KYGRDDVTDEMERARERANAYDFIMEPPNKNFTLVGEKGAOMSGGOKORIAARALVRN 507  
DB 499 KYGRSATQADIERKARANACHDFTIRLPKGIYDVGEGKAQISGGOKORIAARALVRQ 558  
QY 508 PKLILDEATSAIDSEKSAVOALE 533  
DB 559 POVLIDEATSAIDPTESEKRVQSALE 584

Search completed: April 22, 2002, 14:37:44  
Job time: 296 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:35:47 ; Search time 42.7 Seconds  
(without alignments)  
965.116 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739  
Sequence: 1 MIIIGILASLVNGACLPMLP.....SESKSAVQALEKDPTRYSR 541

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.68:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1566.5	57.2	1276	1 DVMS1	multidrug resistin
2	1555.5	56.8	1280	1 DVH01	multidrug resistin
3	1548	56.5	1277	2 JH0502	P-glycoprotein - r
4	1542	56.3	1276	2 A34786	multidrug resistin
5	1540.5	56.2	1276	1 DVHY1C	multidrug resistin
6	1523	55.6	1287	2 S55692	multidrug resistin
7	1455.5	53.1	1279	1 DVH03	multidrug resistin
8	1440.5	52.6	1281	2 I48123	P-glycoprotein iso
9	1430.5	52.6	1276	1 DVMS2	multidrug resistin
10	1421.5	51.9	1278	2 S41646	P-glycoprotein - r
11	1277.5	46.6	1104	1 DVMS1A	multidrug resistin
12	1220	44.5	1321	2 T42228	P-glycoprotein sis
13	1220	44.5	1321	2 T42842	bile salt transpor
14	1134.5	41.4	1294	2 T19982	hypothetical prote
15	1028	37.5	1275	2 T31073	multidrug resistin
16	1002	36.6	1302	2 A41249	multidrug resistin
17	999.5	36.5	1321	2 T23476	hypothetical prote
18	997.5	36.4	1321	2 S27337	multidrug resistin
19	993.5	36.3	1286	2 T02187	probable ABC trans
20	989.5	36.1	1230	2 E85023	P-glycoprotein
21	978	35.7	1283	2 A47377	multidrug resistin
22	975.5	35.6	1229	2 D85023	P-glycoprotein-11k
23	974.5	35.6	1292	2 T48007	P-glycoprotein hom
24	972.5	35.5	1229	2 T52319	P-glycoprotein-11k
25	962	35.1	1222	2 T14805	hypothetical prote
26	962	35.1	1323	2 H85202	hypothetical prote
27	961.5	35.1	1278	2 E86155	probable ABC trans
28	960	35.0	1408	2 T43261	multidrug resistin
29	954	34.8	1318	2 T21266	hypothetical prote

30	949	34.6	1254	2 S27338	P-glycoprotein C -
31	949	34.6	1266	2 T22094	hypothetical prote
32	948	34.6	1266	2 T22090	hypothetical prote
33	942.5	34.4	1229	2 F86155	probable ABC trans
34	942	34.4	1286	2 A42150	P-glycoprotein ppp
35	933	34.1	1302	2 S30327	multidrug resistin
36	927	33.8	1310	2 S30328	multidrug resistin
37	925	33.8	1362	2 T41534	leptomycin B resis
38	917	33.5	1245	2 G66404	probable P-glycop
39	908.5	33.2	1302	2 B41249	multidrug resistin
40	904.5	33.0	1341	2 J60166	LAMP1 protein - L
41	900	32.9	1158	2 T25082	hypothetical prote
42	897.5	32.8	1307	2 T30882	multidrug resistin
43	897	32.7	655	1 DVHY2C	multidrug resistin
44	893	32.6	1270	2 T21265	hypothetical prote
45	892	32.6	1327	2 T21268	hypothetical prote

#### ALIGNMENTS

RESULT 1  
DVMS1  
multidrug resistance protein 1 - mouse  
N:Alternate names: P-glycoprotein 1  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
C:Accession: A33719; A25057; I57510  
R:Raymond, M.; Gros, P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 6488-6492, 1989  
A:Title: Mammalian multidrug-resistance gene: correlation of exon organization with s  
A:Reference number: A33719; MUID:89367274  
A:Accession: A33719  
A:Molecule type: DNA  
A:Residues: 1-1276 <RAY>  
R:Gros, P.; Croop, J.; Housman, D.  
Cell 47, 371-380, 1986  
A:Title: Mammalian multidrug resistance gene: complete cDNA sequence indicates strong  
A:Reference number: A25057; MUID:87028229  
A:Accession: A25057  
A:Molecule type: mRNA  
A:Residues: 1-1276 <GRO>  
A:Cross-references: GB:M4757; MID:g199100; PIDN:AAA79005.1; PID:g387426  
R:Raymond, M.; Gros, P.  
Mol. Cell. Biol. 10, 6036-6040, 1990  
A:Title: Cell-specific activity of cis-acting regulatory elements in the promoter of  
A:Reference number: I57510; MUID:91042535  
A:Accession: I57510  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-21 <RES>  
A:Cross-references: GB:M60348; MID:g199102; PIDN:AAA39513.1; PID:g554199  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
turally and functionally unrelated lipophilic anticancer drugs.  
C:Genes:  
A:Gene: mdrl (pgp1)  
A:Introns: 21/2; 38/3; 92/1; 111/2; 175/2; 233/3; 274/2; 332/3; 370/3; 407/3; 449/3;  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane  
F:1-637 653-1276/Region: duplication  
F:409-603/Domain: ATP-binding cassette homology <ABC1>  
F:426-433/Region: nucleotide-binding motif A (P-loop)  
F:550-554/Region: nucleotide-binding motif B  
F:1051-1247/Domain: ATP-binding cassette homology <ABC2>  
F:1068-1075/Region: nucleotide-binding motif A (P-loop)  
F:1194-1198/Region: nucleotide-binding motif B  
F:773,91,96,103/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:432/Binding site: ATP (Lys) #status predicted  
F:1074/Binding site: ATP (Lys) #status predicted

Query Match 57.2% Score 1566.5; DB 1; Length 1276;  
Best Local Similarity 54.9%; Pred. No. 7.3e-106;





F:1192-1196/Region: nucleotide-binding motif B  
 F:429/Binding site: ATP (lys) #status predicted  
 F:1072/Binding site: ATP (lys) #status predicted

Query Match 56.3%; Score 1542; DB 2; Length 1276;  
 Best Local Similarity 53.8%; Pred. No. 4.4e-104;  
 Matches 299; Conservative 103; Mismatches 98; Indels 56; Gaps 5;

```

QY 1 MIIGTILASVNGACLPMPYLVGEMSDNLISGCLVQTN-----T 39
DB 50 MLVGTLLAIIHGAVALPLMLVFGMDTSPASVGNVSKNSNMSEADKRAMFALEEMTT 109
QY 40 YSEFRLTYVYGIVGVALIFGYIQISLMTITTAAROTKRIKQFHSVLAODIGWEDSCDI 99
DB 110 YAY-----YTTGAGAVLIYAYIQVSWCLAGROHKIKQKFFHAIIMNOEIGMEDVHDV 164
QY 100 GELNTRMT-DIDKISDGIKIALLFQNMSTPSIGLAVLGKMKLTVLTLSTPLIMAS 158
DB 165 GELNTRLTLDVSKINIGIKGFMFOAMATFEGCFITGFTGKLTVLTLSTPLVGLS 224
QY 159 AACSRRVILSTSKELSAVSKAGVAEEVLSTRTVIAFRAQEKELQRSFLNTRRYAMF 218
DB 225 AGIWAKLISFTDKELHAYAKAGVAEEVLAARTVIAFGQKKELE----- 271
QY 219 YFPQWLSCVLAQVRYQNLKDAKDFGIRKRTIASKVSILGAVYFPMNGTYGLAFWYGTSLI 278
DB 272 -----RYNNNLEAKRIGIKKAITANISMGAFLIYASVALAFWYGTSLV 317
QY 279 LINGEPGYTIGTVIAVFFSVIHSYICGAAVPHFETFAIARGAAPHFOVIDKKPSIDNFS 338
DB 318 ISKE--YSIGQVLTVEFVSILGAFVSGQASPNTIEAFANAGAAVEVFKIIDNKPSIDSFS 375
QY 339 TAGKKEPSIGTVVERKVSFVSPRSIKITKGLNLIKSGEMTALVGLNCGSTVYOL 398
DB 376 KSGKHPNIGQNEEFKNHFSYSPKREVOILKGLNLKVSQVYALVNSCGSTVYOL 435
QY 399 LQRLYDDDDGFIWDEDIDALNARHYRDHIGVVSQEPVLFGTTISNNIKYGRDVTDEE 458
DB 436 MQRIYDPLDGVNSIDGDIRINRYLREIIGVVSQEPVLFGTTIANIRGRDVTDEE 495
QY 459 MERAARANAYDTIMEPPKFNILVSGKAGOMSGQKORAIARALVNRKILILDEATS 518
DB 496 IEKAVKCANAYDFTMKLPHQFDPLVGERGAQLSGQKORAIARALVNRKILILDEATS 555
QY 519 ALDESEKSAVOALEX 534
DB 556 ALDESEAVVOALEX 571

RESULT 5
DWHYIC
multidrug resistance protein 1 - Chinese hamster
N:Alternate names: P-glycoprotein pgp1
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 31-Dec-1990 #sequence revision 30-Sep-1992 #text change 19-Jan-2001
C:Accession: A38696; B38696; A27126; S33768; I52823
R:Devine, S.E.; Hussain, A.; Davidge, J.P.; Melera, P.W.
J. Biol. Chem. 266, 4545-4555, 1991
A:Title: Full length and alternatively spliced pgp-1 transcripts in multidrug-resistant
A:Reference number: A38696; MUID:91154265
A:Accession: A38696
A:Molecule type: mRNA
A:Residues: 1-1276 <DEV>
A:Cross-references: GB:M59253; NID:g191154; PIDN:AAA37004.1; PID:g191155
A:Accession: C38696
A:Molecule type: mRNA
A:Residues: 108-1276 <DEL>
A:Cross-references: GB:M59254; NID:g191156; PIDN:AAA37005.1; PID:g191157
A:Experimental source: clone ADX185
A:Accession: B38696
A:Molecule type: mRNA
A:Residues: 1-32, 771-1276 <DE2>

```

```

A:Cross-references: GB:M59252; NID:g191152; PIDN:AAA37003.1; PID:g191153
A:Experimental source: clone ADX124
R:Endicott, J.A.; Juranka, P.F.; Sarangi, F.; Gerlach, J.H.; Deuchars, K.L.; Ling, V.
Mol. Cell. Biol. 7, 4075-4081, 1987
A:Title: Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chinese
A:Reference number: A27126; MUID:88122132
A:Accession: A27126
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 706-1276 <END>
A:Cross-references: GB:M17897; NID:g191158; PIDN:AAA37006.1; PID:g191159
R:Zastawny, R.L.; Ling, V.
Biochim. Biophys. Acta 1173, 303-313, 1993
A:Title: Structural and functional analysis of 5' flanking and intron 1 sequences of
A:Reference number: S33768; MUID:93305724
A:Accession: S33768
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-21 <ZAS>
A:Cross-references: EMBL:L03286
R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.
Cell Growth Differ. 2, 429-437, 1991
A:Title: Analysis of the Chinese hamster P-glycoprotein/multidrug resistance gene pgp
A:Reference number: I52823; MUID:92088970
A:Accession: I52823
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: GB:S81975; NID:9240862
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c
structurally and functionally unrelated lipophilic antitumor drugs.
C:Gene: pgp1
A:Superfamily: multidrug resistance protein: ATP-binding cassette homology
C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding;
F:407-601/Domain: ATP-binding cassette homology <ABC>
F:424-431/Region: nucleotide-binding motif A (P-loop)
F:548-552/Region: nucleotide-binding motif B
F:1050-1246/Domain: ATP-binding cassette homology <ABC2>
F:1067-1074/Region: nucleotide-binding motif A (P-loop)
F:1193-1197/Region: nucleotide-binding motif B
F:87,91,96/Binding site: carbonylate (asn) (covalent) #status predicted
F:430/Binding site: ATP (lys) #status predicted
F:1073/Binding site: ATP (lys) #status predicted

Query Match 56.2%; Score 1540.5; DB 1; Length 1276;
Best Local Similarity 54.0%; Pred. No. 5.7e-104;
Matches 301; Conservative 100; Mismatches 99; Indels 57; Gaps 5;

QY 1 MIIGTILASVNGACLPMPYLVGEMSDNLISGCLVQTN-----T 38
DB 50 MLVGTLLAIIHGAVALPLMLVFGMDTSPASVGNIPINAINNATQVNASIDFGKLEEMT 109
QY 39 YSEFRLTYVYGIVGVALIFGYIQISLMTITTAAROTKRIKQFHSVLAODIGWEDSCDI 98
DB 110 TYAA-----YTTGAGAVLIYAYIQVSWCLAGROHKIKROKFFHAIIMNOEIGMEDVHD 164
QY 99 IGEINTMT-DIDKISDGIKIALLFQNMSTPSIGLAVLGKMKLTVLTLSTPLIMA 157
DB 165 VGEINTRLTLDVSKINIGIKGFMFOAMATFEGCFITGFTGKLTVLTLSTPLVGL 224
QY 158 SAACSRRVILSTSKELSAVSKAGVAEEVLSTRTVIAFRAQEKELQRSFLNTRRYAM 217
DB 225 SAGTWAKLISFTDKELHAYAKAGVAEEVLAARTVIAFGQKKELE----- 272
QY 218 YFPQWLSCVLAQVRYQNLKDAKDFGIRKRTIASKVSILGAVYFPMNGTYGLAFWYGTSL 277
DB 273 -----RYNNNLEAKRIGIKKAITANISMGAFLIYASVALAFWYGTSL 317
QY 278 LINGEPGYTIGTVIAVFFSVIHSYICGAAVPHFETFAIARGAAPHFOVIDKKPSIDNFS 337
DB 318 VISKE--YSIGQVLTVEFVAVLIAFISIGQASPNTIEAFANAGAAVEVFKIIDNKPSIDSF 375

```















GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:34:53 ; Search time 31.47 Seconds  
(without alignments)  
386.834 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739  
Sequence: 1 MIIIGILASIVNGACPLMPL.....SEKSAVQALMKPTPRYSF 541

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCRTGS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1552.5	56.7	1275	3	US-09-120-513-2
2	1552.5	56.7	1275	4	US-09-450-105-2
3	1552.5	56.7	1280	2	US-08-752-447-2
4	1544.5	56.4	1280	2	US-08-583-276-19
5	1544.5	56.4	1280	6	5206352-4
6	1530	55.9	1279	2	US-08-784-649A-2
7	960	35.0	1408	1	US-08-612-521-2
8	935.5	34.2	1349	2	US-08-612-734B-2
9	913.5	33.4	1334	4	US-08-996-545-2
10	913.5	33.4	1334	4	US-09-328-320-2
11	897.5	32.8	1307	1	US-08-395-246C-2
12	740	27.0	1308	2	US-08-996-644-2
13	740	27.0	1308	2	US-09-352-552-2
14	646	23.6	791	4	US-08-394-880B-2
15	627.5	22.9	686	4	US-09-061-764A-15
16	626	22.9	1302	1	US-08-232-537-2
17	621.5	22.7	653	4	US-09-061-764A-2
18	609	22.2	327	1	US-08-463-092B-9
19	609	22.2	327	1	US-08-460-907B-9
20	574.5	21.0	748	4	US-09-061-764A-19
21	496	18.1	707	3	US-08-772-270A-4
22	483	17.6	694	2	US-08-895-522-4
23	483	17.6	694	3	US-09-195-391-4
24	483	16.4	711	3	US-08-772-270A-12
25	449	16.4	747	3	US-08-895-522-1
26	449	16.4	747	3	US-09-195-391-1
27	439.5	16.0	694	2	US-08-895-522-3

28	439.5	16.0	694	3	US-09-195-391-3	Sequence 3, Appl
29	397	14.5	110	1	US-08-466-886-22	Sequence 22, Appl
30	397	14.5	110	4	US-08-466-617-22	Sequence 20, Appl
31	392	14.3	110	1	US-08-466-886-20	Sequence 20, Appl
32	392	14.3	110	4	US-08-466-617-20	Sequence 20, Appl
33	381.5	13.9	109	4	US-08-466-886-24	Sequence 24, Appl
34	381.5	13.9	109	4	US-08-463-092B-6	Sequence 24, Appl
35	357.5	13.1	1528	2	US-08-463-092B-6	Sequence 6, Appl
36	357.5	13.1	1528	2	US-08-462-109A-6	Sequence 6, Appl
37	357.5	13.1	1528	2	US-08-460-907B-6	Sequence 6, Appl
38	357.5	13.1	1528	3	US-08-463-179A-6	Sequence 6, Appl
39	357.5	13.1	1528	3	US-08-461-384B-6	Sequence 6, Appl
40	350.5	12.8	1477	3	US-08-492-459-10	Sequence 10, Appl
41	350.5	12.8	1477	3	US-08-423-753-10	Sequence 10, Appl
42	350.5	12.8	1477	3	US-08-945-994-3	Sequence 3, Appl
43	350.5	12.8	1477	4	US-08-716-873-24	Sequence 24, Appl
44	350.5	12.8	1477	4	US-09-368-431-24	Sequence 24, Appl
45	350.5	12.8	1531	1	US-08-141-893-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-120-513-2  
; Sequence 2, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Ellens, Harna  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDRLB2 AND  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY:  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,513  
; FILING DATE: 22-JUL-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: King, William T  
; REGISTRATION NUMBER: 30,954  
; REFERENCE/DOCKET NUMBER: GP50008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-270-5015  
; TELEFAX: 610-270-5090  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1275 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-120-513-2  
Query Match 56.7%; Score 1552.5; DB 3; Length 1275;  
Best Local Similarity 55.6%; Pred. No. 2.8e-144;

Query Match	56.7%	Score 1552.5	DB 4	Length 1275
Best Local Similarity	55.6%	Pred. No. 2.8e-144		
Matches 308	Conservative 92	Mismatches 105	Indels 49	Gaps 5

RESULT 3  
US-08-752-447-2  
Sequence 2, Application US/08752447  
Patent No. 5894088  
GENERAL INFORMATION:  
APPLICANT: Mechtner, Eugene  
APPLICANT: Koninson, Igor B  
TITLE OF INVENTION: Methods and Reagents for Preparing and  
TITLE OF INVENTION: Using Immunoligal Agents Specific for P-glycoprotein  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESS: McDonnell Boehnen Hulbert & Berghoff Ltd.  
STREET: 300 South Wacker Drive, Seventh Floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752.447  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5994088nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 95,1121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001

TELEFAX: 312-913-9808  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1280 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-752-447-2

Query Match 56.7%; Score 1552.5; DB 2; Length 1280;  
 Best Local Similarity 55.1%; Pred. No. 2.9e-144;  
 Matches 305; Conservative 101; Mismatches 99; Indels 49; Gaps 6;

```

QY 1 MILGLASLVGACPLMPVLGEMSD-----NLSGCLVOT--NTSFF----- 43
DB 51 MVTGLTAIIGHAGPLMLVFGEMTDIFANAGNLEDMSTNTNSDINDTGFFNNLEED 110
QY 44 --RLTVYVGIGVALIFGYIQISLMTTARQTKRIRKOPFHSVLAODIGMFDSCDICE 101
DB 111 MTRYAYYSIGAGVLAAYIQVSWCLAGROHKKIRKOPFHAIMREIGMFDVHDVGE 170
QY 102 LNTMT--DIDKISDGIKIALFPQNMSTFISGLAVGLVKGLTVTLSTSPILMASAA 160
DB 171 LNTRLTDVSKINEYIGDKIGMFQSMATFFTGFTGKWLTVLTAISPVLGLSAA 230
QY 161 ACSRMVLSLTKELSAKAGAAVEVLSSTRVIAFPAQEKELORSFLNITRAMYTF 220
DB 231 VWAKILSFDEKELAAKAGAAVEVLAIRTVIAFGQKKELE----- 275
QY 221 POWLSCVLXEVRYTONLKDADFGIKRTIASKVSGLAVYFPMNGTGYLAFTYGTSLIN 280
DB 276 -----RYKNKLEAKRIGIKKAITANISIGAAFLITIASYALAAFTYGTSLIN 323
QY 281 GEPGYTIGTVLAVFESVISHSYCIGAAVPHETFAIARGAAFHIFQVIDKRPISDNFSTA 340
DB 324 GE--YSIGQVLTVFESVILGAFSVQASPSEAFANARGAAVEIFKIIDNKPISDYSKS 381
QY 341 GYKPESTIGYEFKRVSNYPSRSIKILKGLNLRIRKSGEYVALVGLSGSKSTYVQLQ 400
DB 382 GHKPDNIGKNEFRNVHVSYSRSREVKILKGLNLRKVSQGVYALVGLSGSKSTYVQLQ 441
QY 401 RLYPDDGFIVWDENDIALNVRHYRDIHIGVSOEPLYFGTTISNNIKYGRDDVTDEME 460
DB 442 RLYPDTGMSVSDODITRTINRFLREITIGVSOEPLYFATTIANIRYGNBNVTMDIE 501
QY 461 RAAREANAYDEIMEFPNKNTLVGEKQAGMGGOKORIAIARALVRNPKILLDEATSAL 520
DB 502 KAVEANAYDEIMKLPHKFDTLVGERGAQLSGGOKORIAIARALVRNPKILLDEATSAL 561
QY 521 DSEKSAVQAALEK 534
DB 562 DTESEAVVQVALDK 575

```

RESULT 4  
 US-08-583-276-19  
 Sequence 19, Application US/08583276  
 Patent No. 5837536  
 GENERAL INFORMATION:  
 APPLICANT: McDonagh, Kevin T.  
 APPLICANT: Nienhuis, Arthur  
 APPLICANT: Tolstoshev, Paul  
 TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
 TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Carella, Byrne, Bain, Giffillan,  
 ADDRESSEE: Cecchi & Stewart  
 STREET: 6 Becker Farm Road  
 CITY: Roseland  
 STATE: New Jersey

COUNTRY: USA  
 ZIP: 07068  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch diskette  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: DM4.V2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/583,276  
 FILING DATE: 05-JAN-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/332,444  
 FILING DATE: 31-OCT-1994  
 APPLICATION NUMBER: 07/887,712  
 FILING DATE: 22-MAY-1992  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1280 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE:  
 DESCRIPTION: protein  
 US-08-583-276-19

Query Match 56.4%; Score 1544.5; DB 2; Length 1280;  
 Best Local Similarity 55.1%; Pred. No. 1.8e-143;  
 Matches 305; Conservative 99; Mismatches 101; Indels 49; Gaps 6;

```

QY 1 MILGLASLVGACPLMPVLGEMSD-----NLSGCLVOT--NTSFF----- 43
DB 51 MVTGLTAIIGHAGPLMLVFGEMTDIFANAGNLEDMSTNTNSDINDTGFFNNLEED 110
QY 44 --RLTVYVGIGVALIFGYIQISLMTTARQTKRIRKOPFHSVLAODIGMFDSCDICE 101
DB 111 MTRYAYYSIGAGVLAAYIQVSWCLAGROHKKIRKOPFHAIMREIGMFDVHDVGE 170
QY 102 LNTMT--DIDKISDGIKIALFPQNMSTFISGLAVGLVKGLTVTLSTSPILMASAA 160
DB 171 LNTRLTDVSKINEYIGDKIGMFQSMATFFTGFTGKWLTVLTAISPVLGLSAA 230
QY 161 ACSRMVLSLTKELSAKAGAAVEVLSSTRVIAFPAQEKELORSFLNITRAMYTF 220
DB 231 VWAKILSFDEKELAAKAGAAVEVLAIRTVIAFGQKKELE----- 275
QY 221 POWLSCVLXEVRYTONLKDADFGIKRTIASKVSGLAVYFPMNGTGYLAFTYGTSLIN 280
DB 276 -----RYKNKLEAKRIGIKKAITANISIGAAFLITIASYALAAFTYGTSLIN 323
QY 281 GEPGYTIGTVLAVFESVISHSYCIGAAVPHETFAIARGAAFHIFQVIDKRPISDNFSTA 340
DB 324 GE--YSIGQVLTVFESVILGAFSVQASPSEAFANARGAAVEIFKIIDNKPISDYSKS 381
QY 341 GYKPESTIGYEFKRVSNYPSRSIKILKGLNLRIRKSGEYVALVGLSGSKSTYVQLQ 400
DB 382 GHKPDNIGKNEFRNVHVSYSRSREVKILKGLNLRKVSQGVYALVGLSGSKSTYVQLQ 441
QY 401 RLYPDDGFIVWDENDIALNVRHYRDIHIGVSOEPLYFGTTISNNIKYGRDDVTDEME 460
DB 442 RLYPDTGMSVSDODITRTINRFLREITIGVSOEPLYFATTIANIRYGNBNVTMDIE 501
QY 461 RAAREANAYDEIMEFPNKNTLVGEKQAGMGGOKORIAIARALVRNPKILLDEATSAL 520
DB 502 KAVEANAYDEIMKLPHKFDTLVGERGAQLSGGOKORIAIARALVRNPKILLDEATSAL 561
QY 521 DSEKSAVQAALEK 534
DB 562 DTESEAVVQVALDK 575

```

RESULT 5

5206352-4  
 Patent No. 5206352  
 APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman, Michael M.  
 TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
 NUMBER OF SEQUENCES: 4  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/622,836  
 FILING DATE: 24-SEP-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 892,575  
 FILING DATE: 01-AUG-1986  
 APPLICATION NUMBER: 845,610  
 FILING DATE: 28-MAR-1986  
 SEQ ID NO:4:  
 LENGTH: 1280  
 5206352-4

Query Match 56.4%; Score 1544.5; DB 6; Length 1280;  
 Best Local Similarity 55.1%; Pred. No. 1.8e-143;  
 Matches 305; Conservative 99; Mismatches 101; Indels 49; Gaps 6;

QY 1 MILGILASLVNACPLPLMPLVIGEMSD-----NLISGCLVOT--NTYSFF----- 43  
 DB 51 MVVGTALAIHAGLPLMLVGEEMTDIFANAGNLEDMSNTNRSDINDTGFNMLEED 110  
 QY 44 --RLTLVYVIGVAAALIFGIQISLMTITPAROTKRIKOFPHSVLAODIGWDSDDIGE 101  
 DB 111 MTRVAYYSGIGVLYAAYIOVSFMCAGROHRIKROFHAIKROELGEMDVHDVGE 170  
 QY 102 LMTBMT-DIDKISDGIIDGKIALLFONKSTFSGIANGVKGKMLTLVTSTPLIASAA 160  
 DB 171 LNTRLDDVSKINQVIGDKIGMFQSMATFFTGFIYGTGRGKMLTLVLAISVGLISAA 230  
 QY 161 ACSRWVLSLTKELSAVSKAGAAVEELSSIRTVIAFRAQEKELQSFLLNTRYAMEFY 220  
 DB 231 VMKILISFTDKELLYAAGAAVEELAIRVIAFGGKKLE----- 275  
 QY 221 POWLSVCLFVRYTONLKDADGKIRTIASVSLGAVYFFNNGYGLAFWYGTSLILN 280  
 DB 276 -----RYNNLEAKRIGIKKAITANISIGAFLIYASALAFWYGTSLVLS 323  
 QY 281 GEPGYTIGVLAFFSVIHSYCIGAAPHFEFPATARGAAPHFOYIDKPSIDNESTA 340  
 DB 324 GE--YSIGVLYVFSVLIGAFVSGQSPSIEFAFANARGAAYEIFKIDKPSIDSTKS 381  
 QY 341 GYKPESTIEGVEKNSFNPSRPSIKILKGLMRITKSGETVALVNGSGKSTVQLOQ 400  
 DB 382 GHRPDNIKGNLEFRNVHFSYPSRKEVKILKGLMKVOSGGTVALVNGSGCKSTTVQLOQ 441  
 QY 401 RLTPDDGFIWDENDIRALNVHRHDIGVSOEPEVLFEGTTISNNIKYGRDVTDEME 460  
 DB 442 RLTPDEGAVSVGDGDIRTNVFRLEIIGVVSQEPVLFATTAENIRYGERENVTDEIE 501  
 QY 461 RAAREANAYDFIMEPKNFTLVGEKGAOMSGOKORAIARALVNRPKLLIDEATSA 520  
 DB 502 KAKENANADFIKLPKEDTLVGERGAKORAIARALVNRPKLLIDEATSA 561  
 QY 521 DSEKSAVQAALKE 534  
 DB 562 DTSEAVVQVALDK 575

RESULT 6  
 US-08-784-649A-2  
 Sequence 2, Application US/08784649A  
 Patent No. 5830697  
 GENERAL INFORMATION:  
 APPLICANT: SLIC, Braundmtr I  
 APPLICANT: Chen, Gang  
 TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO

TITLE OF INVENTION: CYCLOSPORIN MODULATION  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson  
 STREET: 2200 Sand Hill Road  
 CITY: Menlo Park  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94025  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/784,649A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sherwood, Pamela J  
 REGISTRATION NUMBER: Reg. No. 5830697 36,677  
 REFERENCE/DOCKET NUMBER: 06037/007001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-322-5070  
 TELEFAX: 415-854-0875  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1279 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-784-649A-2

Query Match 55.9%; Score 1530; DB 2; Length 1279;  
 Best Local Similarity 54.9%; Pred. No. 4.8e-142;  
 Matches 304; Conservative 100; Mismatches 100; Indels 50; Gaps 7;

QY 1 MILGILASLVNACPLPLMPLVIGEMSD-----NLISGCLVOT--NTYSFF----- 43  
 DB 51 MVVGTALAIHAGLPLMLVGEEMTDIFANAGNLEDMSNTNRSDINDTGFNMLEED 110  
 QY 44 --RLTLVYVIGVAAALIFGIQISLMTITPAROTKRIKOFPHSVLAODIGWDSDDIGE 101  
 DB 111 MTRVAYYSGIGVLYAAYIOVSFMCAGROHRIKROFHAIKROELGEMDVHDVGE 170  
 QY 102 LMTBMT-DIDKISDGIIDGKIALLFONKSTFSGIANGVKGKMLTLVTSTPLIASAA 160  
 DB 171 LNTRLDDVSKINQVIGDKIGMFQSMATFFTGFIYGTGRGKMLTLVLAISVGLISAA 230  
 QY 161 ACSRWVLSLTKELSAVSKAGAAVEELSSIRTVIAFRAQEKELQSFLLNTRYAMEFY 220  
 DB 231 VMKILISFTDKELLYAAGAAVEELAIRVIAFGGKKLE----- 275  
 QY 221 POWLSVCLFVRYTONLKDADGKIRTIASVSLGAVYFFNNGYGLAFWYGTSLILN 280  
 DB 276 -----RYNNLEAKRIGIKKAITANISIGAFLIYASALAFWYGTSLVLS 323  
 QY 281 GEPGYTIGVLAFFSVIHSYCIGAAPHFEFPATARGAAPHFOYIDKPSIDNESTA 340  
 DB 324 GE--YSIGVLYVFSVLIGAFVSGQSPSIEFAFANARGAAYEIFKIDKPSIDSTKS 380  
 QY 341 GYKPESTIEGVEKNSFNPSRPSIKILKGLMRITKSGETVALVNGSGKSTVQLOQ 400  
 DB 381 GHRPDNIKGNLEFRNVHFSYPSRKEVKILKGLMKVOSGGTVALVNGSGCKSTTVQLOQ 440  
 QY 401 RLTPDDGFIWDENDIRALNVHRHDIGVSOEPEVLFEGTTISNNIKYGRDVTDEME 460  
 DB 441 RLTPDEGAVSVGDGDIRTNVFRLEIIGVVSQEPVLFATTAENIRYGERENVTDEIE 500  
 QY 461 RAAREANAYDFIMEPKNFTLVGEKGAOMSGOKORAIARALVNRPKLLIDEATSA 520

DB 501 KAVKRNANVDFIMKLPKHFEDLVGERGAOLSGGOKORAIARALVNPRIILLDEATSAL 560  
QY 521 DESKSAVOALEK 534  
DB 561 DIESEAVOVALDK 574

RESULT 7  
US-08-612-521-2

; Sequence 2, Application US/08612521  
; Patent No. 5786463  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B  
; APPLICANT: Skatrud, Paul L  
; APPLICANT: Thornevell, Susan J  
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Division/AEH  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,521  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hamilton, Amy E  
; REGISTRATION NUMBER: 33,894  
; REFERENCE/DOCKET NUMBER: X-9693  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3169  
; TELEFAX: 317-276-1294  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1408 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-612-521-2

Query Match 35.0%; Score 960; DB 1; Length 1408;  
Best Local Similarity 37.1%; Pred. No. 1.3e-85;  
Matches 210; Conservative 112; Mismatches 182; Indels 62; Gaps 8;  
QY 1 MILGILASLVNGACLPMLPLVLGEMSDNLSGCLVOTNTYSEFFRLT----- 46  
DB 156 MYLGLVLAAGSCQPLMLTIFGRITSTNTYAVI-ANOISGGCLTPEPSALQAAKDL 214  
QY 47 -----LYYVGIVAAIFGYIOWISMTITPAQTKRIRKOFHFSVLAODIGWEDSCD 98  
DB 215 KTQSGHNALYLAIGMFLATLWYFIMWVTGELNSKRIRERYLAVALRQELAVYDDLG 274  
QY 99 IGEIWTIRM-TDIDKISDGDGDKIALLFQNMSTFSGIAGLVKGMKLTLYTSTSLINA 157  
DB 275 AEEVATRIOTDCHLVQEGSEKVALVFOYAGTFVCGEVLAFVRSPLAGALVSLPVIWL 334  
QY 158 SAAASRWYISLTKSELAYSKAGAAVEVLSIRTVIAFRAOEKELQSLNITRYAM 217  
DB 335 CGGIMWTMAAKGTALDIAAKAGSLAEVIGSIRYQAF-GKEKTLGKRF----- 384  
QY 218 FYFPMOLLSCVLAFVRYTONLKDANDEGIKRTIASKVSLGAAVYFENNGTYGLAFWYGTSL 277  
DB 385 -----ADHIEQSKIYGRKGSIFEGFGISIMFEVLYAAYALAFYFGIL 427

QY 278 ILNGEPGYTIGTVLAFVFSVSHSSYCIGAAYPHFETALARGAAFIHFOVINDKPSINDF 337  
DB 428 VSNQGA--DSGIYINVFMSILIGFSMALAPELAAVTARGAALKFTITRVAIDSA 485  
QY 338 STAGYKPESEIEGTVEFKNVSPNPSRSIKILGNLRKSGEVALVNGSGKSTVQO 397  
DB 486 SEEGFKPDGLRGEISPEENVKFPYPSRPSIPIIKGFTTTEAGKTFALVAGSSGSKSTVVS 545  
QY 398 LIQRLYDPDDGFIWDENDIRALNVRHYRDIHGVYSQEVYLEGTTISNNIKTG-----RD 452  
DB 546 LIERFYDPVSGYVKLDGRDIRSLNLMWLRQOIGLVQSQEPITLGTIVRGVNEHGLGSRYE 605  
QY 453 DYTDEE---MERAREANAAYDEIMEFPKFTVLGEGKAQSGGOKORAIARALVRNP 508  
DB 606 NASLEKEFELVKACVDAHANFIMKLPQGYDTWGERGMLLSGGOKORAVARALVSDP 665  
QY 509 KILLIDEATSALDESKSAVOALEK 534  
DB 666 RILLIDEATSALDTQSEGIYODALDK 691

## RESULT 8

US-08-612-734B-2

; Sequence 2, Application US/08612734B  
; Patent No. 5914246  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B  
; APPLICANT: Skatrud, Paul L  
; APPLICANT: Robin, Matthew B  
; TITLE OF INVENTION: Multiple Drug Resistance Gene of  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center, DC1501  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.A.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/612,734B  
; FILING DATE: 08-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Craig, Anne I.  
; REGISTRATION NUMBER: 32,976  
; REFERENCE/DOCKET NUMBER: X-9681  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-354-9570  
; TELEFAX: 617-354-4043  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1349 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-612-734B-2

Query Match 34.2%; Score 935.5; DB 2; Length 1349;  
Best Local Similarity 37.0%; Pred. No. 3.1e-83;  
Matches 203; Conservative 110; Mismatches 191; Indels 45; Gaps 7;  
QY 1 MILGILASLVNGACLPMLPLVLGEMSDNLSGCLVOTNTYSEFF--RLT---LYYVGIVAA 55  
DB 111 IIVSAICATAAGAAALPLFLITFGSLA-SAFQGISICLTMPIYHEFYHKLTVNVLAFYVYLGIA 169



```

OY 56 ALIEGYIQLSWITTAARQKRIKQFHSVLADIGMFDSCDIGELNTRMT-DIDKISD 114
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 EFTVYVSTVGFTGHEHLQKIRENYLEALIQNMAYFPDKLGEVTRITADTNLID 229
OY 115 GIDGKALLFQNMSTFSGIAGLVKWKMLTVLSTPLMASAACSRMVISLTKEL 174
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 AISEKGLTLTAPATFTAPITAVYKWKLLALICTSTIVALVWVGSGSRIVYKRSKI 289
OY 175 SAYSKAGAAVEVLSTRTVIAFRAQEKELORSFLLNITRYAMFFQWMLSCVLFVRY 234
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 290 ESTGAGGTVAEEVISTIRNATFECTDKLAK-----QY 322
OY 235 TQNLKADKFGIKRTIASKVSLGAVFFPMNGTYGLAFWYGTSLILNCEPGYITGLAVF 294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 323 ETHLAEEKMGVKKQOVLIGMMIGMFGIMFNSVGLGFMGSGRFVYKRE--VWQVQVLTVL 380
OY 295 FSVYHSYCIGAAPHFETRIANGAAPHFOVDKRPISDNFSTAGYKPSISGYEKF 354
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 381 MSILIGFSIGNAVAPNQAFTNGVAAAKIYSTIDRSPLDPYSDDEKVLDFEGNTEFR 440
OY 355 NVSPNPSRSIKILKGLNLRKISGETVALVGLNGSGKSTVYVOLLQRLYPDDEGIWDE 414
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 NVKIIYSRPVYMEVYSLSMRPAKKTALVGPSSGSKSTVYGLVERFYLPVGQVLLDG 500
OY 415 NDIRALNVRRHDIHGVYSGEPVLFGTISNNIKY-----RDVYDEMEKARE 465
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 501 HDIGTLMLRMLRQOISLVSGEPVLFSTTIFRNIEHLIGTKFEHESKDKIRELVENARM 560
OY 466 ANADFTIMEFPNKNTLVGEKAGOMSGGOKRIALAVRNPKILLDEFTSLDSEK 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 561 ANADFTIMEPEGYDVTNGGKFLSGGOKRIALAVRIVDPKILLDEFTSLDSEK 620
OY 526 SAVQAALER 534
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 621 GYVQAALDK 629

```

RESULT 9  
 US-08-996-545-2  
 : Sequence 2, Application US/08996545  
 : Patient No. 5928898  
 : GENERAL INFORMATION:  
 : APPLICANT: Skatrud, Paul L.  
 : APPLICANT: de Waard, Maarten A.  
 : APPLICANT: Peery, Robert B.  
 : APPLICANT: Andrade, Alan C.  
 : TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
 : TITLE OF INVENTION: Aspergillus nidulans  
 : NUMBER OF SEQUENCES: 3  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Eli Lilly and Company  
 : STREET: Lilly Corporate Center  
 : CITY: Indianapolis  
 : STATE: Indiana  
 : COUNTRY: U.S.  
 : ZIP: 46285  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/996,545  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Webster, Thomas D.  
 : REGISTRATION NUMBER: 39,872  
 : REFERENCE/DOCKET NUMBER: X-11766  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 317-276-3334  
 : TELEFAX: 317-276-2763  
 : INFORMATION FOR SEQ ID NO: 2:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1334 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-996-545-2

```

Query Match 33.4%; Score 913.5; DB 2; Length 1334;  
 Best Local Similarity 39.0%; Pred. No. 4.5e-81;  
 Matches 194; Conservative 96; Mismatches 169; Indels 39; Gaps 4;

```

OY 47 LYYGVGVAALIFGYIDISMLITTAARQKRIKQFHSVLADIGMFDSCDIGELNTRM 106
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 153 LFEVYLIGFEVYVSTVGFTGHEHLQKIRENYLEALIQNMAYFPDKLGEVTRIT 212
OY 107 T-DIDKISDGIQKALLFQNMSTFSGIAGLVKWKMLTVLSTPLMASAACSRM 165
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 TADTNLIQDISEKVGTLTALATFTVTAFTAVYKWKLLALICTSTIVALVWVGSG 272
OY 166 VISLTKELSAYSKAGAAVEVLSTRTVIAFRAQEKELORSFLLNITRYAMFFQWML 225
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 IIRKSKSLDSYGAGGTVAEEVISTIRNATFECTDKLAK----- 312
OY 226 SCVLAEXVRYQNLKADKFGIKRTIASKVSLGAVFFPMNGTYGLAFWYGTSLILN 265
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 -----QYEVHLEADEAKWKGNQIVMGFMIGAMFGLMYSNYGLGFMGSGRFVLDG--AV 363
OY 286 TIGTVLAVFVSIVHSYICGAAPHFETRIANGAAPHFOVDKRPISDNFSTAGYKPE 345
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 DVGDLTLVLAAILGSSSLGVSPNAQFTNVAATAAKITGTDKQSPLDYRSEGTLD 423
OY 346 SIEGTEFEKNVSPNPSRSIKILKGLNLRKISGETVALVGLNGSGKSTVYVOLLQRL 405
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 HFEGHIELRNKHIYPSRPVYMEVYSLSMRPAKKTALVGPSSGSKSTVYGLVERFY 483
OY 406 DDGFTWDENDIRALNVRRHDIHGVYSGEPVLFGTISNNIKY-----RDVYD 456
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 484 VRGTVLLDGHDKIDLNRMLRQOISLVSGEPVLFSTTIFRNIEHLIGTKFEHESK 543
OY 457 EEMERAREANAYDFTIMEFPNKNTLVGEKAGOMSGGOKRIALAVRNPKILLDEA 516
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 544 ELINAAANAHADFTIRALPEGYETNGGKFLSGGOKRIALAVRIVDPKILLDEA 603
OY 517 TSALDSEKSAVQAALER 534
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 604 TSAIDTKSEGVQAALER 621

```

RESULT 10  
 US-09-328-320-2  
 : Sequence 2, Application US/09328320  
 : Patient No. 6228615  
 : GENERAL INFORMATION:  
 : APPLICANT: Skatrud, Paul L.  
 : APPLICANT: de Waard, Maarten A.  
 : APPLICANT: Peery, Robert B.  
 : APPLICANT: Andrade, Alan C.  
 : TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
 : TITLE OF INVENTION: Aspergillus nidulans  
 : NUMBER OF SEQUENCES: 3  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Eli Lilly and Company  
 : STREET: Lilly Corporate Center  
 : CITY: Indianapolis  
 : STATE: Indiana  
 : COUNTRY: U.S.  
 : ZIP: 46285  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/328,320  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/996,545  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Webster, Thomas D.  
;; REGISTRATION NUMBER: 39,872  
;; REFERENCE/DOCKET NUMBER: X-11766  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 317-276-3334  
;; TELEFAX: 317-276-2763  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1334 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-09-328-320-2

Query Match 33.4%; Score 913.5; DB 4; Length 1334;  
Best Local Similarity 39.0%; Pred. No. 4.5e-81;  
Matches 194; Conservative 96; Mismatches 169; Indels 39; Gaps 4;

QY 47 LYYGIGVAAALIFGTIQTSLMTIITAAROTKRIKOFHSHVLAODIGWDFSCDIGELNTRM 106  
D 153 LFFVYLGIGVEVTVTVGVTVGFIYGEHATOKIRREYLYESILRONIGYDKLGAGEVTTRI 212  
QY 107 T-DIDKISDGIKIDKIALLFQNMSTFSGIAGLVKGMKLTIVTLSTSLIMASAAACSRM 165  
D 213 TADNLNLDGISEKVGGLTALMAFFVTAFITAYVKKWKLALICSTIVALTVMGGSQF 272  
QY 166 VISLTSKELSAVSKAGAAVEEVLSSIRTVIAFRAQEKELQSFLLNITRYAMFEPQWL 225  
D 273 IIXKSKSLDSYGAGGYAAEEVSIIRNATAFGTODKAK----- 312  
QY 226 SCVLAFFRYTONLKADDFGKTRITASKVSLGAVYFFNNGTYGLAFWYGTSLILNGEGY 285  
D 313 -----QYEVHLDEAEKMGCTKNOIVMGFMIGMGLYSNVLGFWMSRFLVDG--AV 363  
QY 286 TIGTVLAVFESVHSSYICGAAPHFEEFATARGAAPHIQQVIDKPSIDNFSTAGRYPE 345  
D 364 DVGILTVLWMLIGTSLGNVSPNAQAFITVAVAANAIFGTIDRQSPLDDEYSNEGKTL 423  
QY 346 SIEGTVEKKNYSFNYSPRSIKILKGLNLRIKSETVALVGLNGSKSTVVQLQRLYDP 405  
D 424 HFECHIELRNKHIYPSRPETVMEVDSLSMPACKRTALVGPSSGSKSTVVGIVLERFTMP 483  
QY 406 DDGFTWDENDIRALNVRHYDHIGVVSQEPVLFCTTISNNIKYG-----RDVYTD 456  
D 484 VRGVLLDGHDIKIDKLNLMWLRQOISLVSQEPVLFCTTITKKNIRHGLIGTKYENESD 543  
QY 457 EEMRAAREAAAYDIFEMFPMKFTNIVGEEKGAKSGGOKORAIARALVRPKIILDEA 516  
D 544 ELIENAAAMAAHDPITALPGYEITNGRGFLSLGGOKORAIARAVASDPKILLDEA 603  
QY 517 TSALDESSEKSAVQAAL 534  
D 604 TSALDTKSEGVQAALER 621

RESULT 11  
US-08-395-246C-2  
Sequence 2, Application US/08395246C  
Patent No. 5773214  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF ASPERGILLUS FLAVUS

;; NUMBER OF SEQUENCES: 4  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Eli Lilly and Company  
;; STREET: Lilly Corporate Center  
;; CITY: Indianapolis  
;; STATE: IN  
;; COUNTRY: USA  
;; ZIP: 46285  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/395,246C  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Plant, Thomas G.  
;; REGISTRATION NUMBER: 35784  
;; REFERENCE/DOCKET NUMBER: X9683  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 317-276-2459  
;; TELEFAX: 317-277-1917  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1307 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-395-246C-2

Query Match 32.8%; Score 897.5; DB 1; Length 1307;  
Best Local Similarity 36.0%; Pred. No. 1.7e-79;  
Matches 196; Conservative 109; Mismatches 194; Indels 45; Gaps 8;

QY 6 LASLVNCAELPLMPVLVGEKMDNLISGLVQNTNTYSF-----RLTIYVGIGVAAALIFG 60  
D 67 LAALAGGALPLFLVFLGRLT-STFQDIATHRIITYDHHHELTKNVYFIVYLGAAEFVAI 125  
QY 61 YIQSLMTIITAARQKRIKOFHSHVLAODIGWDFSCDIGELNTRM-TDIDKISDGIK 119  
D 126 YIATVGFITTDHYVQQRVEYFOAILKONIAFFDTLGAGETTRITADNLNLDGISEK 185  
QY 120 IALLFQNMSTFSGIAGLVKGMKLTIVTLSTSLIMASAAACSRMVISLTSKELSAVSK 179  
D 186 VGLALTLGSTVTVTAFLIYIKNMKALICSASLALLTLTMGGCSTLMILFSKMALEYGR 245  
QY 180 AGAAVEEVLSSIRTVIAFRAQEKELQSFLLNITRYAMFEPQWLSCVLAFFRYTONLK 239  
D 246 GASMAEDITLDSIRVAAANAOE-----TLAR-----KESHLK 278  
QY 240 DAKDGIKTRTASKSLGAVYFFNNGTYGLAFWYGTSLILNGEPEYTGTVLAVFESYIH 299  
D 279 DAEQGMKSKYIPALMVALLCINTLNGLGFWMSRFLVEGISIRKAGDVLTIMAITL 338  
QY 300 SSYICGAAPHFEEFATARGAAPHIQQVIDKPSIDNFSTAGYRPESIEGTVEERKNVSFN 359  
D 339 GSYNLGNVAPNGQALSDVAANAASKLYGTIDRQSPLDALSDGKTLFEYRGVNIQNIHV 398  
QY 360 YPSRPSIKILKGLNLRIKSETVALVGLNGSKSTVVQLQRLDPPDGFIMDENDIR 419  
D 399 YPSRPEVTVADLSQYIPAGKTTAFVGPSSGSKSTIISLBERFYDPVAGTILMDGHDLOT 458  
QY 420 LNVHRHYDHIGVVSQEPVLFCTTISNNIKYG-----RDVYTD-----EEMRAAREAAAYD 470  
D 459 LNLKRLQOMSLVSOEPLFTTAENTIRYGLISRFKESTYEIRKVEAAARAANAHD 518  
QY 471 FIMEFPMKFTNIVGEEKGAKSGGOKORAIARALVRPKIILDEATSALDESSEKSAVQA 530  
D 519 FIMALPNQYDINI--ESRSLSGGOKORAIARAIKIDPKILLDEATSDLTCKSEKLVQA 576



Db 261 EAGHETRLQI-----HAQANSFAEGITLACVKAHVAFGRND----- 296  
Qy 219 YEPOMLLSCVLXEVRYTONLKDADKGIRKTASKVSLGAVYFPNMGTGLAFWYGTSLI 278  
Db 297 -----SLVKKRFDEY---LVEAHKVKKISPLGLLFSAEVITLYLGAFMOGIMHF 346  
Qy 279 LNGEGYITIGVLAVFVSIIHSSYICGAAPHFEFATARGAAPHIIOYIDKKPSIDMS 338  
Db 347 GRGEIG-TAGDIFTVLSVSLASINLTLAPSTIEFSRAASAQAFRLIDESSINPYG 405  
Qy 339 TAGYKPESEIEGVEKKNVSPNPSRPSIKILKGLNLRKSGETVALVGLNGSGKSTVYL 398  
Db 406 KEGLEPERVLGDVELENTFESYTPRGITVLDNFSKVPAGKVTALVQSSGSKSTIYGL 465  
Qy 399 LQRLVDPDDGFLMVDENDIRALNVHRHDHIGVVSQEPVLFGTTISNNIKYG-----RD 452  
Db 466 LERWNPSTSGAIRLDGNLISELNVGMLRRNVLYVQOEPVLFQGSYFDMIRYGLVGTPMEN 525  
Qy 453 DVTDEMER-----AAREANAYDIFIMEFPKNTVLVEKGAOMSGGOKRIATARALVRMPK 509  
Db 526 ASREQMERVOEAPKLAIAHEFISLTDGYDPLIGERGGLSGGOKORVALARAVSQPK 585  
Qy 510 ILILDEATSALDESSESAYQAALAK 534  
Db 586 VLLDEATSALDPHAETIVQAKALDK 610

RESULT 14  
US-08-394-880B-2  
; Sequence 2, Application US/08394880B  
; Patent No. 5705352  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene Of  
; TITLE OF INVENTION: Aspergillus Fumigatus  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company/Patent Division  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: US  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,880B  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant G., Thomas  
; REGISTRATION NUMBER: 35784  
; REFERENCE/DOCKET NUMBER: X-9682  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 276-2459  
; TELEFAX: (317) 277-1917  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 791 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-394-880B-2

Query Match 23.6%; Score 646; DB 1; Length 791;  
Best Local Similarity 31.1%; Pred. No. 5,4e-55;  
Matches 170; Conservative 97; Mismatches 223; Indels 56; Gaps 11;

Qy 1 MTLGLTSLVNGACPLRPLVILGEMSDNLISGLVQNTNTYSFRRLTY-----VGIGV 54  
Db 181 LALAFLLVSSGITMSITFSIGKIMDTSTYKATTEGNELEGLSLPMFYGALAGILTLTGA 240  
Qy 55 AALIGYIQLISLMTITTAAROTRIKOPFHSVLAODICWFDSCDGLGELNTRMTD-----ID 110  
Db 241 AA---NYGRITILTRVGERIVARLSKLFQRFVQDAEFEDNRRGDLISRLSSPTIYG 297  
Qy 111 K-ISDGIQDKTALLFQNNSTYSIGLAVGLVKGWKLTLVTLSPLIMASAAACSRMVISL 169  
Db 298 KSIQNLSDGLAAAVSAGAFGLMAYVSL---KLSSILALLPLPIGLGAFYFGAIRNL 353  
Qy 170 T---SKELSAVSKAGAAVAEVLSSIRTYIAFPAQKEKLQSRSLNITRAMYFFQWML 226  
Db 354 SNOIQRNIGTLTK---IAERLGNVKTSSQSFGE----- 384  
Qy 227 CVLXEVRYTONLKDADKGIRKTASKVSLGAVYFPNMGTGLAFWYGTSLILNGEPCYT 286  
Db 365 -VLEVRRTNNVQKRIEFLGKKSLSATFFSSTGFAGNMTILALLYVGGMVQSG--ATT 441  
Qy 287 IGTVAVEFVSIIHSSYICGAAPHFEFATARGAAPHIIOYIDKKPSIDNFSFAGYKPPES 346  
Db 442 IGLTSLFMYTAYAGSSMFGSSFSYSELMKKGVAASRLFEILDROPTIS--PTKGEKVAS 499  
Qy 347 IEGYVEFKNVSPNPSRPSIKILKGLNLRKSGETVALVGLNGSGKSTVYOLLQRLYPD 406  
Db 500 ARGPIREFNVETSYTPRAVPIFRDLNFEIPQGTVAIVGPSGGKRSITASTILLFYSPT 559  
Qy 407 DGFIVDENDIRALNVHRHDHIGVVSQEPVLFGTISNNIKYGRDDVTDEEMERAREA 466  
Db 560 EGRVILIGDKDITHMAKSLRKIGIVSQEPVLFSGTIAENIYKGRPOAKRSIVAAAKRA 619  
Qy 467 NAYDIFIMEFPKNTVLVEKGAOMSGGOKRIATARALVRNPKILILDEATSALDESSES 526  
Db 620 NC-QETSPDDLDIQVGRGALSGGOKRIATARALIKDDPILILDEATSALDAESPT 678  
Qy 527 AVQAAL 532  
Db 679 LVNSAL 684

RESULT 15  
US-09-061-764A-15  
; Sequence 15, Application US/09061764A  
; Patent No. 6284879  
; GENERAL INFORMATION:  
; APPLICANT: Faustman, Denise L.  
; TITLE OF INVENTION: TRANSPORT ASSOCIATED PROTEIN SPLICE VARIANTS  
; TITLE OF INVENTION: AND MODEL FOR IMMUNE DIVERSITY  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Yankwich & Associates  
; STREET: 130 Bishop Allen Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: United States of America  
; ZIP: 02139  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: Microsoft Word 97 SR-1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/061,764A  
; FILING DATE: April 16, 1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Yankwich, Leon R  
; REGISTRATION NUMBER: 30,237

REFERENCE/DOCKET NUMBER: MGH-002.0 US/MHG-1247.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-491-4343  
TELEFAX: 617-491-8801  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 686 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-061-764A-15

Query Match 22.9%; Score 627.5; DB 4; Length 686;  
Best local similarity 31.7%; Pred. No. 2,9e-53;

Matches 177; Conservative 102; Mismatches 176; Indels 103; Gaps 15;

QY 15 LPLMP-----LVIGEMSDNLI---SGCLVQT-----NTYSF---PRLTLVYVIGVYAL 57  
DB 149 LPLVAAFFFLVAVLGETLIPHYSGRVIDILGDFDPHAFASAFEMCLFSFGSSISAG 208  
QY 58 IFGYIQISLIIITAAQTKRIKQFPHSVLAODIGWFDSCDIDELNTRM--TDIDKISDGI 116  
DB 209 CRG---GCTYMSRINLRIRQLESSLRQDGFQETKGTGLNSRLSDTTLMSMWL 264  
QY 117 GDKIALLFQNMSTFISGLAVLKGW-----KLTIVTSTSPILMASAACSRLVYSLT 170  
DB 265 P-----LNAVLLRSLVKVVGLY--GFMLSTSPRLTLLSLHMFPTIAEKVYNTRHQEV 318  
QY 171 SKELSAVSKAGAAEVLSSIRTVIAFRAOE-----KELOKSFLLN 211  
DB 319 REIQDAVARAGQVREAVGIGLQTVRSFGAEHEVCRYKEAKEQCRQLYWRDLERALYLL 378  
QY 212 ITRYAMFYFQWLLSCVLXVRYTONLKDADGIRKTIASKVSLGAVYFPMGTYGIAF 271  
DB 379 IRRVHLGVOMLMSGL-----QOMOD-----GELTQGSLSFM----- 413  
QY 272 WYGTSLINCEPGYITGTVLAVFFSVYHSSYCIGAAPHETFAIARGAAFIPOVIDKK 331  
DB 414 -----IYQSVGSYVOTLVYIIGDMLSNV-----GAAEKVESYMDRQ 450  
QY 332 PSIDNESTAGYKPESTIEGVEFNKVSFNPSPRSIKILKGLNLRKSGETVALVGLNGSG 391  
DB 451 PNLPSPTTLA--PTTLQGVVKFQDVSEFAYPNRPDPVILKGLTFLRPGEVYALVGPNGSG 508  
QY 392 KSTVYVOLLQRLYDPDGFIVDENDIRALNVRHNRDHIGVVSQEPVLEGTISNNIKYR 451  
DB 509 KSTVYVALLQNLQPTGGVLLDEKPIQYEHCHYLSQVSVSGDEPVLFGSVRRNNIAYGL 568  
QY 452 DDVYDEEMERAAEANAAYDFIMEFPNKFNTLVGEKGAQMSGGQKORTAIAATVARNPKIL 511  
DB 569 QSCEDDKVMAAQAHAHDDFIQEMEHGIYTDVGEKGSQIAAGQKORLAIAATVARDPVL 628  
QY 512 ILDEATSLDSEKSAVO 529  
DB 629 ILDEATSLDQCEQALQ 646

Search completed: April 22, 2002, 14:34:58  
Job time: 130 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:34:15 ; Search time 61.37 Seconds

(without alignments)  
652.984 Million cell updates/sec

Title: US-09-873-409-7

Perfect score: 2739  
Sequence: 1 MIIIGILASLVNGACILPMLP.....SESKSAVQALAEKDTPTYSF 541

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_1101.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1566.5	57.2	1276	14	AA35199
2	1566.5	57.2	1276	14	AA35199
3	1561.5	57.0	1281	22	AAE00310
4	1557.5	56.9	1281	22	AAE00308
5	1557.5	56.9	1281	22	AAE00309
6	1557.5	56.8	1283	22	AA81065
7	1555.5	56.8	1280	18	AAW4073
8	1555.5	56.8	1280	21	AAV58186
9	1555.5	56.8	1280	22	AA81959
10	1552.5	56.7	1275	21	AAV78879
11	1552.5	56.7	1275	22	AA860409

12	1552.5	56.7	1280	19	AAW48997
13	1552.5	56.7	1280	19	AAW48999
14	1552.5	56.7	1280	22	AA81068
15	1552.5	56.7	1280	22	AAE00304
16	1552.5	56.7	1281	22	AAE00303
17	1550.5	56.6	1280	22	AAU04347
18	1548.5	56.5	1280	22	AA81064
19	1546.5	56.5	1280	15	AA863624
20	1546.5	56.5	1280	19	AAW48998
21	1546.5	56.5	1280	19	AAW48999
22	1546.5	56.5	1280	21	AAV58187
23	1546.5	56.5	1280	22	AA81066
24	1546.5	56.5	1280	22	AAE00306
25	1544.5	56.4	1280	8	AA70452
26	1543.5	56.4	1280	11	AA804868
27	1543	56.3	1272	21	AAV70597
28	1541	56.3	1272	21	AAV70596
29	1540.5	56.2	1280	14	AAW44297
30	1539	56.2	1279	22	AA81067
31	1539	56.2	1279	22	AAE00307
32	1536	56.1	1276	22	AAV58189
33	1530	55.9	1279	19	AAW80294
34	1528	37.5	1275	20	AAW82594
35	960	35.0	1408	19	AAW62871
36	942	34.4	1243	21	AAW39102
37	942	34.4	1286	21	AAW39101
38	935.5	34.2	1349	20	AAV16434
39	923.5	33.7	1205	21	AAW39103
40	913.5	33.4	1334	20	AAV02630
41	905	33.0	1254	21	AAV92173
42	897.5	32.8	1307	17	AAW92255
43	865.5	31.6	1333	22	AAW62495
44	835.5	30.5	1261	22	AAW62493
45	792	28.9	1248	22	AAW62488

#### ALIGNMENTS

RESULT	ID	AA35199	standard	Protein	1276	AA
AC	AA35199					
DT	28-JUL-1993	(first entry)				
DE	Mouse multidrug resistance protein.					
KW	mdr gene; Lambda DR11 clone.					
XX	Mus musculus.					
OS	US5198344-A.					
PN	30-MAR-1993.					
XX	15-JUL-1986;	86US-0885951.				
XX	15-JUL-1986;	86US-0885951.				
XX	06-FEB-1991;	91US-0652311.				
XX	(MASI ) MASSACHUSETTS INST TECHNOLOGY.					
XX	Croop JM, Gros P, Housman DE;					
XX	WPI: 1993-126077/15.					
XX	N-PDB: AAQ38950.					
XX	DNA sequence which confers multi-drug resistance on sensitive mammalian cells - used to preserve bone marrow cells during chemotherapy to prevent infection					

PS Disclosure; Fig 8; 22pp; English.

XX The lambda DR11 mdr gene was isolated as described in AA038950.

CC The amino acid sequence deduced from its open reading frame has

CC potential glycosylation sites and is likely to be a membrane

CC protein. The encoded protein appears to contain (starting at the

CC 5'-end) three hydrophobic domains, a neutral domain, three

CC additional hydrophobic domains and a second neutral domain. There

CC are ATP-binding sites within the neutral domains. It is

CC hypothesised that increased synthesis of the membrane protein may

CC alter membrane physiology to cause decreased intracellular drug

CC accumulation.

XX

XX

SQ Sequence 1276 AA;

Query Match 57.2%; Score 1566.5; DB 14; Length 1276;

Best Local Similarity 54.9%; Pred. No. 7.8e-144;

Matches 307; Conservative 99; Mismatches 94; Indels 59; Gaps 6;

QY 1 MITGLIATLVNGACLPMPVLVIGESDNL-----SGCLVQTN----- 39

DB 50 mlglfllaingltlplmlvfgmtsfkkaasilpstngsgpnstllisslslee 109

QY 40 --YSEFRLTYVYGIVGVALIFGYIQISLWITPARQTKRIKOFHVSVALAODIGWFD 96

DB 110 maiyay-----yfgigsgvllvaylqyslwcslaagrqhkrqkffhaimgelgwfdv 164

QY 97 CDIGELNTRMT-DIDKISDGIQDKTALLFQNNSTFISGLAVGKWKLTIVTSTPLI 155

DB 165 hdvgeIntrldvskIndigdkigmftgsIttlaagflligfsgwKltIvIavsp11 224

QY 156 MASAACSRMTISLTSKELSAVSKAGAVAEVLSIRTVIAFRAQEKLOSFLLNTRY 215

DB 225 glssalwakvltstfnkclgayaakagavaeevlaairvtaifsgqkele----- 274

QY 216 AMFYEPQWLSCVLXEVRYTONLKDADKFGIKRTIAKSVSLGAVYFPMNGTYGLAFWYGT 275

DB 275 -----TynknlleaKnvgIkKaltasIsIqIayllvyaSalaIafwygt 317

QY 276 SLIINGEGYTIIGVLAIVFVSIVHSYICGAANVPHFETPARAGAAFIPOVDDKKPSID 335

DB 318 slvlsne--ysIgevlvIstfllgltfslgIhlapleaanaargaafIkIldpsId 375

QY 336 NFSTAGYKESIEGTEVERKSNFNPSPRSIKILKGLMRKSGFTVALVINGSGKSTV 395

DB 376 sfstkgkyprdsImgnlekvnhfypsrsevgllKqIlnlkKsgqvtIvngsgcgkstt 435

QY 396 VQLIQRLVDPDDGFLMDENDIRALNVRHRDHIQVSOEPLYEGTTISNNIKYGRDVT 455

DB 436 vqlmgrlvdpdgyvsvIdqgdlrtlnvrylrelIqvgvsgepvlftfaenlrygredvt 495

QY 456 DEEMERARERANADTFMEFNKNTLVGEGAGMSGQKQRIALARKLVNPKLLIDE 515

DB 496 mdeIekavkeanaydfImklphqfdltlvgergaqIsgqkqrlaIaralvnpkIId 555

QY 516 ATSAIDSEKSAVOALEK 534

DB 556 atsaldeseavvaalok 574

RESULT 2

AAV58188 ID AAV58188 standard; Protein; 1276 AA.

XX AAV58188;

XX

DT 14-MAR-2000 (first entry)

XX

DE Murine multidrug resistance-1 (MDR-1).

XX

KW Multidrug resistance; MDR-1; P-glycoprotein;

KW transmembrane efflux pump; haematopoietic stem cell; transduction;

KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;

KW gene therapy; gene replacement; genetic defect; thalassemia;

KW Gaucher's disease; sickle cell anaemia; leukemia; ex vivo expansion;

KW cytokine.

OS Mus sp.

XX

XX W09961589-A2.

XX

XX

XX 02-DEC-1999.

XX

XX 27-MAY-1999; 99WO-US11825.

XX

XX 28-MAY-1998; 98US-0086988.

XX

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX

XX Sorrentino B, Bunting K;

XX

XX WPI; 2000-072615/06.

DR N-PSDB; AA249334.

DR

XX

PT Ex vivo expansion of hematopoietic stem cells transduced with a

PT sequence encoding human multidrug resistance-1, used for bone marrow

PT transplantation -

PS Disclosure; Page 93-101; 113pp; English.

XX

CC This sequence represents murine multidrug resistance

CC protein MDR-1. MDR-1 is a transmembrane efflux pump,

CC responsible for the export of drugs from cells, particularly

CC cancer cells. The invention relates to transducing hematopoietic

CC stem cells with nucleic acid encoding an MDR protein

CC and culturing the modified cells. The modified hematopoietic stem

CC cells are useful in bone marrow transplantation (to reconstitute

CC hematopoietic systems in patients who have undergone chemotherapy or

CC radiation therapy) and in ex vivo gene therapy of genetic defects in

CC cells derived from hematopoietic stem cells, e.g., thalassemia,

CC Gaucher's disease, sickle cell anemia or leukemia. The modified

CC cells can also be used to identify factors involved in regulating

CC proliferation and differentiation in hematopoietic stem cells.

CC Haematopoietic stem cells that express MDR-1 will be protected against

CC chemotherapy agents, so can be engrafted while the patient is

CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells

CC provides sufficient cells to permit standard biochemical analysis.

CC Overexpression of MDR-1 allows cytokine-driven expansion of

CC haematopoietic stem cells by at least 10-fold compared with a maximum

CC of 4-fold in known procedures.

XX

SQ Sequence 1276 AA;

Query Match 57.2%; Score 1566.5; DB 21; Length 1276;

Best Local Similarity 54.9%; Pred. No. 7.8e-144;

Matches 307; Conservative 99; Mismatches 94; Indels 59; Gaps 6;

QY 1 MITGLIATLVNGACLPMPVLVIGESDNL-----SGCLVQTN----- 39

DB 50 mlglfllaingltlplmlvfgmtsfkkaasilpstngsgpnstllisslslee 109

QY 40 --YSEFRLTYVYGIVGVALIFGYIQISLWITPARQTKRIKOFHVSVALAODIGWFD 96

DB 110 maiyay-----yfgigsgvllvaylqyslwcslaagrqhkrqkffhaimgelgwfdv 164

QY 97 CDIGELNTRMT-DIDKISDGIQDKTALLFQNNSTFISGLAVGKWKLTIVTSTPLI 155

DB 165 hdvgeIntrldvskIndigdkigmftgsIttlaagflligfsgwKltIvIavsp11 224

QY 156 MASAACSRMTISLTSKELSAVSKAGAVAEVLSIRTVIAFRAQEKLOSFLLNTRY 215

DB 225 glssalwakvltstfnkclgayaakagavaeevlaairvtaifsgqkele----- 274

QY 216 AMFYEPQWLSCVLXEVRYTONLKDADKFGIKRTIAKSVSLGAVYFPMNGTYGLAFWYGT 275





```

XX  Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KM  MDR1; drug bioavailability; transgenic animal; genetic model.
XX
OS  Canis familiaris.
FH  Key
FT  Location/Qualifiers
FT  Misc-difference 197
XX  /note= "His of Genotypec substituted by Gln"
XX  WO200123540-A2.
PD  05-APR-2001.
XX
PF  28-SEP-2000; 2000WO-US26767.
XX
PR  28-SEP-1999; 99US-0156510.
XX  (GENT-) GENTEST CORP.
XX
PI  Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX  WPI: 2001-235373/24.
DR  N-PSDB; AAD03504.
XX
PT  New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT  for determining the bioavailability of drugs and for screening for dog
PT  PGP inhibitors.
XX
PS  Claim 17; Page 91-93; 111pp; English.
XX
CC  The invention relates to dog P-glycoprotein (PGP) also referred
CC  as multidrug transporter (MDR1) and nucleic acids encoding them.
CC  The invention also includes fragments and biologically functional
CC  variants of dog P-glycoprotein. PGP and their nucleic acids are
CC  useful for determining the bioavailability of drugs and for
CC  screening PGP inhibitors. They are useful for the diagnosis and
CC  treatment of conditions characterised by PGP activity, by
CC  reducing or increasing PGP activity in a cell. PGP nucleic acids
CC  are used as oligonucleotide probes. Complements of PGP nucleic
CC  acids are useful as antisense oligonucleotides, to induce a PGP
CC  'knockout' phenotype. They are used to prepare a non-human
CC  transgenic animal, which are valuable as genetic models for
CC  human diseases.
CC  The present sequence is dog P-glycoprotein (PGP) allelic variant.
CC  This sequence is also referred as Genotype A protein. The PGP
CC  enzyme functions as an efflux pump exporting small molecules
CC  across the cell membrane. This enzyme is a member of the ABC
CC  transporter family.
XX
SQ  Sequence 1281 AA;

```

```

Query Match 56.9%; Score 1557.5; DB 22; Length 1281;
Best Local Similarity 54.5%; Pred. No. 6e-143;
Matches 303; Conservative 98; Mismatches 104; Indels 51; Gaps 5;

```

```

OY  1 MLGLIASLVNACLPMPVLVGEISDNLTISGCL-----VQNTNYSFPR--- 44
DB  51 mlvgmaailhgaalplmnlvfgnmdsfanagisrnkflpylinestnntnqhfinhle 110
OY  45 -----LTIYVVGIGVAALIFGYIOISLWITTAARQTKRRKQFFHSVLAQDQGWPDSCI 99
DB  111 eemtyayyysgigavvaaylqysfwclaaagrgllkirkdfihaimqelgwtdvvhv 170
OY  100 GEINRMT-DIKISDGDIGDIALFLQNNSTFSIGLAVGLVKGWKLTVLSTSPILMAS 158
DB  171 glntrcttdvskinegjdkgmftgfatftgftvgftgwkltvllaispylgls 230
OY  159 AAGSRMTVISTSKELSAVSKAGAAVEEVLISIRVIAFRAQEKELQRFSLNITRYAMF 218
DB  231 aaiwakllsfcdkellayakagavaeevlaairvtatfgqgkhele----- 277

```

```

OY  219 YEPQWLLSCVLNFEVRYTONLKDADFGIKRTIASKVSLGAVVEFMNGTYGLAFWYGTSLI 278
DB  278 -----zynkleeakgiklkatltanlsigaftllyasyalafwygtsliv 323
OY  279 LMGEGYTTIGTVLAAVFFSVTHSSYTCGAADVPEFPAIRGAAPHIPOVIDKKPSIDNFS 338
DB  324 lse--yslgvlytvlfsylgafsigaspslaefanargaaeyefkldkpsidsys 381
OY  339 TAGYRPESIEGVEEPKNSVFNPSPRSIKITLGLNLRIRKSGFTVALVGLNSGKSTVQL 398
DB  382 ksgkhpndikgnlefknvhnfsyprskvckilqglnlkvsgqgtvalvngsgckstvtql 441
OY  399 LQRLYPDDEFTMDVENDIRALNVRHRYDHIGVSGEPVLFQGTTSNNIKYGRDVTDEE 458
DB  442 mgrlpyrdigmwcidgqddrtlnvrhlreilgtvsgqepvlftlaenlyrgrenvtnde 501
OY  459 MERARERANAYPTMEFPKENTLYGEGKAGQSGQKQKIALARALYRPRKLLIIDETS 518
DB  502 lekavkeanaydfimkplkftdlvgergaqlsgqgkqrlataraalvmpkllldeats 561
OY  519 ALDSEKSAVOALEK 534
DB  562 aldteseavvqvaldk 577

```

```

RESULT 5
AAE00309
ID  AAE00309 standard; Protein: 1281 AA.
XX
AC  AAE00309;
XX
DT  13-JUN-2001 (first entry)
XX
DE  Dog P-glycoprotein (PGP) allelic variant (Genotype B).
XX
KW  Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW  MDR1; drug bioavailability; transgenic animal; genetic model.
XX
OS  Canis familiaris.
FH  Key
FT  Location/Qualifiers
FT  Misc-difference 25
FT  /note= "Asn of Genotypec substituted by Lys"
FT  Misc-difference 197
FT  /note= "His of Genotypec substituted by Gln"
XX
PS  WO200123540-A2.
XX
PD  05-APR-2001.
XX
PF  28-SEP-2000; 2000WO-US26767.
XX
PR  28-SEP-1999; 99US-0156510.
XX  (GENT-) GENTEST CORP.
XX
PI  Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX  WPI: 2001-235373/24.
DR  N-PSDB; AAD03505.
XX
PT  New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT  for determining the bioavailability of drugs and for screening for dog
PT  PGP inhibitors.
XX
PS  Claim 17; Page 99-102; 111pp; English.
XX
CC  The invention relates to dog P-glycoprotein (PGP) also referred
CC  as multidrug transporter (MDR1) and nucleic acids encoding them.
CC  The invention also includes fragments and biologically functional
CC  variants of dog P-glycoprotein. PGP and their nucleic acids are
CC  useful for determining the bioavailability of drugs and for
CC  screening PGP inhibitors. They are useful for the diagnosis and

```

[illegible][illegible]

```

Db      324 vlake--ysigvltvftsvlifaifvgqaspsleaafanargaafefkldnkpsidsy 381
OY      338 STGKYRPESTIEGVEFKNVFNPNRPSIKILKGLNRISGEVALVGLNGSKSTVVO 397
Db      382 sksgnkpdmnlkgnltnvftsyprkevklkylglnlkvgsgqlvalvngsgcgskstvq 441
OY      398 LLOLYDDPDGFTIMVENDIRALNVHRHYPHIGVSOEPLVFTTISNNIKYGRDVTDE 457
Db      442 lmgrllyplegmsvvdqgdrltnvrlfrellygvsgepylfatleentlrygredvtmd 501
OY      458 EMERAAEANADEFIMEFPKFNPTLVGEKGAOMSGGOKRIARALVRNPKIILDEAT 517
Db      502 eiekavneanaydfimklpkkftllygergaqlsggqkqrlataraalvrnplkllldeat 561
OY      518 SALDSEKSAVOALEK 534
Db      562 saldtesaavvqvaldk 578

RESULT 7
AAM44073
AAM44073 standard; protein; 1280 AA.
AC      AAM44073;
XX      26-JUN-1998 (first entry)
XX      Human multidrug resistance P-glycoprotein MDR1.
XX      Human multidrug resistance P-glycoprotein; MDR1, prokaryotic homologue.
XX      Homo sapiens.
FH      Key Location/Qualifiers
FT      Region 1..640
FT      /note= "MDR1-N from Fig 1"
FT      641..1280
FT      /note= "MDR1-C from Fig 1"
XX      MO9740160-A1.
XX      30-OCT-1997.
XX      24-APR-1997; 97MO-N000216.
XX      24-APR-1996; 96EP-0201094.
XX      (UYGR-) RIJXSUNIV GRONINGEN.
XX      PA
XX      PI Bolhuis H, Konings WN, Van Veen HW, Venema K;
XX      DR WPI; 1997-535844/49.
XX      PT Prokaryotic homologue of human multiple drug resistance protein -
XX      PS used to screen for compounds that inhibit, or avoid, drug resistance
XX      Claim 10; Fig 1; 35pp; English.
XX      CC The present invention describes a recombinant or isolated nucleic acid
CC      (1), derived from a prokaryotic gene, which encodes at least a specific
CC      and/or functional part of a transporter protein (TP), or its
CC      derivatives, which has functional and/or structural similarity with the
CC      P-glycoprotein (Pg) encoded by the human multidrug resistance
CC      (MDR1) gene. The present sequence represents the human MDR1 protein,
CC      derived from MDR1-N and MDR1-C as shown in the specification in
CC      figure 1. (1) is used to express recombinant proteins; its fragments
CC      are also useful as probes and primers for detection and amplification
CC      of related DNA. The protein produced, or cells expressing them, are
CC      used to determine if substances can inhibit, or avoid, MDR proteins,
CC      and in a screening method for identifying compounds that inhibit
CC      transport of cytotoxic substances from cells. Also, cells with a
CC      transmembrane protein, especially where expressed from (1), can provide

```

```

CC (additional) MDR, particularly for use as a model system to study
CC mechanisms of action of Pg.
XX
SQ Sequence 1280 AA;
Query Match 56.8%; Score 1555.5; DB 18; Length 1280;
Best Local Similarity 55.2%; Pred. No. 9.4e-143;
Matches 306; Conservative 100; Mismatches 99; Indels 49; Gaps 6;
OY 1 MIIIGTASLVNAGCLPMLPVLGEMSD-----NLISGLVQT--NITYSF----- 43
Db 51 mvyvgtlaailhggaglpimmlvfgemtdifanagnledlmsnltrsdindtgffimnleed 110
OY 44 --RLTLVYGIGVALIFGVYQISLWITTAORFKRIKOPFHSVLNODIGMPCSDIGE 101
Db 111 mtrayayysgigagvivaayiqvswfwaiaargqlnkirkqffhaimgeiygfwhvhyge 170
OY 102 LNTFRMT-DIDKISDGIKDIALLFQNNSTFSGIAGVLVKGWKLTVLTSTPLIMASAA 160
Db 171 Intrltdvskinegigdkigmfqsmatffgfvgtfgrwkltvlilaipylgisaa 230
OY 161 ACSRMVLSLSKELSAISKAGAAVEVLSIRYVIARRAQEKELQRFLLNITRYAFYF 220
Db 231 vvakllsftfdkellayakagavaeevlaalrtvialtggqkkele----- 275
OY 221 PQWLSCVLFVRYTONLKDQKDFGIRKRTIASKVSLGAVFPMNGTGLAFWGTSLILN 280
Db 276 -----tyknblseekrigikaltanaisigaafillyasjalatvgttlvls 323
OY 281 GEPGYTIGTVLAFPSYIHSSYCIGAAPPHFEPFARGAAPHILOYIDKKPSIDNESTA 340
Db 324 ge--ysigvltvftsvlifaifvgqaspsleaafanargaafefkldnkpsidsysks 381
OY 341 GYKPESTIEGVEFKNVFNPNRPSIKILKGLNRISGEVALVGLNGSKSTVVOLO 400
Db 382 ghkpdnlkgnltnvftsyprkevklkylglnlkvgsgqlvalvngsgcgskstvqlmq 441
OY 401 RLTPDDGFTIMVENDIRALNVHRHYPHIGVSOEPLVFTTISNNIKYGRDVTDEME 460
Db 442 rlypdegmsvvdqgdrltnvrlfrellygvsgepylfatleentlrygrentvmade 501
OY 461 RAAREANAYDEFIMEFPKFNPTLVGEKGAOMSGGOKRIARALVRNPKIILDEATSA 520
Db 502 kavkeanaydfimklpkkftllygergaqlsggqkqrlataraalvrnplkllldeatsal 561
OY 521 DSEKSAVOALEK 534
Db 562 dtesaavvqvaldk 575

RESULT 8
AAY58186
ID AAY58186 standard; Protein; 1280 AA.
XX
AC AAY58186;
XX
XX 14-MAR-2000 (first entry)
XX
XX Human wild-type multidrug resistance-1 (MDR-1) protein.
XX
XX Multidrug resistance; MDR-1; P-glycoprotein;
XX transmembrane efflux pump; haematopoietic stem cell; transduction;
XX bone marrow transplantation; chemotherapy; radiation therapy; cancer;
XX gene therapy; gene replacement; genetic defect; thalassaemia;
XX Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
XX cytokine; wild-type.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Msc-difference 185
FT /note= "This residue is Val in a mutant MDR-1

```

FT (AAV58187) \*

XX W09961589-A2.

PN 02-DEC-1999.

PD 27-MAY-1999; 99WO-US11825.

PE 28-MAY-1998; 98US-0086988.

PR (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PA Sorrentino B, Bunting K;

PI WPI: 2000-072615/06.

DR N-PSDB; AAZ49332.

XX Ex vivo expansion of hematopoietic stem cells transduced with a

PT sequence encoding human multidrug resistance-1, used for bone marrow

PT transplantation -

XX Claim 10; Page 71-79; 113pp; English.

PS This sequence represents human wild-type multidrug

CC resistance protein MDR-1. MDR-1 is a transmembrane efflux

CC pump, responsible for the export of drugs from certain

CC cells, particularly cancer cells. Wild-type MDR-1 shows

CC increased resistance to etoposide and decreased resistance to vinca

CC alkaloids compared with a mutant form (AAV58187) where the Gly at

CC position 185 is replaced by Val. The invention relates to transducing

CC hematopoietic stem cells with nucleic acid encoding an MDR protein

CC and culturing the modified cells. The modified hematopoietic stem

CC cells are useful in bone marrow transplantation (to reconstitute

CC hematopoietic systems in patients who have undergone chemotherapy or

CC radiation therapy) and in ex vivo gene therapy of genetic defects in

CC cells derived from hematopoietic stem cells, e.g., thalassemia,

CC Gaucher's disease, sickle cell anemia or leukemia. The modified

CC cells can also be used to identify factors involved in regulating

CC proliferation and differentiation in hematopoietic stem cells.

CC Hematopoietic stem cells that express MDR-1 will be protected against

CC chemotherapeutic agents, so can be engrafted while the patient is

CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells

CC provides sufficient cells to permit standard biochemical analysis.

CC Overexpression of MDR-1 allows cytokine-driven expansion of

CC hematopoietic stem cells by at least 10-fold compared with a maximum

CC of 4-fold in known procedures.

CC XX

XX Sequence 1280 AA:

SO

Query Match 56.8%; Score 1555.5; DB 21; Length 1280;

Best Local Similarity 55.2%; Pred. No. 9.4e-143;

Matches 306; Conservative 100; Mismatches 99; Indels 49; Gaps 6;

1 MITGIASLVGACTLPMPVLIGMSD-----NLISCLVQTF--NTSFF----- 43

DB 51 mvyagtlaaillngaglpimnmlvgemctdfanagnledmshnitrndindgymleed 110

QY 44 --RLTLVVGIGVALIFGYIQISLWITTAARQTKRIKQFHSVLADQIDMFPSCDIGE 101

DB 111 mtryayyyysgagvylvaaylqvswfclaaagrqhkhkqgfhhaimgeiswfdvhdage 170

QY 102 LNTWMT-DIDKISGIGDKIALLFQNMSTESIGLAVGILVKGWKLTVLTSTSPIMASAA 160

DB 171 Intltldvskinegldkqgmftgsmatftgftvgftgwkltlvlaalspylglsaa 230

QY 161 ACSRWVISTSKELSAVSKAGVAEEVLSSTRVIAFAQEKELORSRLNITRWAMYF 220

DB 231 vwaiklsfckellayakagaaveevlaalrtvialtggqkkel----- 275

QY 221 POWLSCVLFVRYTONLDAKDPGIRKRTIASKVSIGAVVFNMGTGLAWFGSTSLIN 280

DB 276 -----rynhkleekrigikkatantansigaalliyasyalatwtygttlvls 323

QY 281 GEPGTTTCTVAVFFSVSHSSYCIGANPHETRAIANGAAPHFOVDKRPIDNSTFA 340

DB 324 ge-yslqgvlvtffsvlilgafswgapsleatanaagaayelfkildmnpdsidsysks 381

QY 341 GYKPESTGTEVPEKFNVSNNYSPRSPIKILKGLNLRKSGEFTVALVGLSGSKSPYVQLO 400

DB 382 ghpknllkgnljetfnvhtyspikvklkglnlkvsgqvalvgnsgcskstvtqimq 441

QY 401 RLVPDDGFIMVENDIRALNVRHYRDHIGVSOEPVLFGTTSNNIKYGRDVTDEME 460

DB 442 rlydptegmavsvdgdtdltlnvrlfrellyvvsqpvlfatlaenlygrenvmele 501

QY 461 RAARANAIDFIMEPPKFNFLVGEKGQMSGGQKQRIARIALVLRNKKILIDEAISAL 520

DB 502 kavkeanaydfimklphkfdclvgergaqlsgqkqrlataraivnphkllldeatsal 561

QY 521 DSESKAVQALEK 534

DB 562 dtesaavvqvaldk 575

## RESULT 9

AAAB81959

ID AAB81959 standard; protein: 1280 AA.

XX AAAB81959;

XX 02-JUL-2001 (first entry)

XX Human MDRL.

XX Human; MDRL: multi-drug resistance; cholesterol uptake;

KW hypercholesterolaemia; hypocholesterolaemia; atherosclerosis;

XX coronary artery disease; cerebral vascular disease.

OS Homo sapiens.

XX W0200121762-A2.

XX 29-MAR-2001.

XX 22-SEP-2000; 2000WO-US26099.

XX 23-SEP-1999; 99US-0155819.

XX (BARN-) BARNES-JEWISH HOSPITAL.

XX Stenson WF, Tessner T;

XX WPI: 2001-328100/34.

XX Modulating cellular cholesterol uptake for treating hyper or

PT hypocholesterolemia, involves administering an agent that inhibits or

PT increases the expression of multiple drug resistance-1 -

XX Example 1; Page 43-47; 47pp; English.

XX The present invention describes a method of modulating cholesterol uptake

CC in cells, involving administering an agent capable of inhibiting the

CC multi-drug resistance protein MDRL. This is useful in the prevention and

CC treatment of cholesterol-related diseases, including

CC hypercholesterolaemia, hypocholesterolaemia, atherosclerosis, coronary

CC artery disease and cerebral vascular diseases. The present sequence is

XX the human MDRL protein.

XX Sequence 1280 AA:

Query Match 56.8%; Score 1555.5; DB 22; Length 1280;

Best Local Similarity 55.2%; Pred. No. 9.4e-143;

Matches 306; Conservative 100; Mismatches 99; Indels 49; Gaps 6;

QY	1	MILSLSLVVGACLPPLPLVLGMSN-----NLISGCLVQT--NNYSEF-----	43
Db	51	mvysltlaaiihngaelplmmvlvfgemtdifanagnlelmsnitchrsindtqffmleed	110
QY	44	--RLTLVVGIGVALALFFGYIQISLWITTAARQTRIRKQFPHSVLAQDIGWDSQDIG	101
Db	111	mtreyayysgagvalvvaayqvsfwcslaagrqhiklrkqffhaimgejfvdhvge	170
QY	102	LNTKMT-DIDKISGIGDKIALLEQNNSTPSTIGLAVLGKWKLTLYLTSTPLMASAA	160
Db	171	lntrltdvskinegldgklygmffgsmatfftglyftrgwkltlylalspvlqisa	230
QY	161	ACSRMVISLSEKELSAVSKAGAVAEVYLSIRYIARQAEKELORSFLLNTRYAFYF	220
Db	221	vwakllsfckellayakagaavaeevlaalrtvlatqgqkke-----	275
QY	221	POWLLSLCLAEVRYTOMLKDKKDGIRKTIAASKVSLGAVVEFNNGTGLAFWYGTSLIN	280
Db	276	-----rynklnseekrigrkikatanisigaefillyasvalafwygtlvls	323
QY	281	GEPCGTICTVLAVFESVYHSSYICIGAAVPHETFAIRGAAPHIIOYIDKPSIDNESTA	340
Db	324	ge--ysiqqvltvfsvlqifasvgsaspsleafanaargaayeikfidnkpidsysks	381
QY	341	GKKPEISGTEVEFNVSNNYSPRSIRIKLGNLRISGENGVAVLGNSGKSNVQVLLQ	400
Db	382	ghkphnklqneltnmvfshysprvevkllylnkvgsgqvalavngsgcgskctvqimq	441
QY	401	RLYPDDDFIIVWENDIRALNWRHYRPHIGVSOEPLYFGTTISNNRYKRGDVTDEME	460
Db	442	rllypdtgmvavsdgddrltnlvrfreljyvgsepylrfatlaenlrygrenvtmdete	501
QY	461	RAARAAAYDPIMEFPKFNLTVGEGKAGMSGGOKRIARALARVNRPKIILDEATSAI	520
Db	502	kavkeaanaydfimkprhfcldvgergaqlsgqkqrfalalaralvnrpkililideatsal	561
QY	521	DSESKAVQALEK 534	
Db	562	dteseavvqvaldk 575	
RESULT 10			
AA78879			
ID	AA78879 standard; Protein: 1275 AA.		
XX	AA78879;		
DT	19-MAY-2000 (first entry)		
XX			
DE	Rat mdrlb2 (multispecific drug transporter) protein sequence.		
XX			
KW	Rat; mdrlb2; multispecific drug transporter; cancer; inflammation;		
KW	cardiovascular disease; central nervous system disorder;		
KW	autoimmune disease; kidney disease; drug formulation.		
XX			
OS	Rattus sp.		
XX			
PN	US6025160-A.		
PD	15-FEB-2000.		
XX			
PF	22-JUL-1998; 98US-0120513.		
XX			
PR	22-JUL-1998; 98US-0120513.		
XX			
PA	(SMIK ) SMITHKLINE BEECHAM CORP.		
XX			
PI	Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;		
XX	WPI; 2000-181810/16.		
DR	N-PSDB; AAZ90198.		
XX			
PT	Isolated polynucleotide encoding a rat mdrlb2 drug transporter		

This sequence represents a rat mdrlb2 multispecific drug transporter protein. The mdrlb2 nucleotide sequence has cytosstatic, antinflammatory, cardiact, neuroprotective, immunosuppressive and nephrotoxic activity. Understanding the functioning of the mdrlb2 polynucleotide and protein in transgenic animal models is useful for treating and preventing diseases such as cancer, inflammation, cardiovascular disease, central nervous system disorders, autoimmune disease, and kidney disease. The use of the protein in cell based, membrane based, or binding assays may enhance drug formulation, selection of formulation excipients and compound design.

Claim 6: Column 17-20; 17pp: English.

polypeptide, useful for treatment of e.g. cancer, autoimmune disease, central nervous system disorders -

Query Match 56.7%; Score 1552.5; DB 21; Length 1275;  
 Best Local Similarity 55.6%; Pred No.1.8e-14;  
 Matches 308; Conservative 92; Mismatches 105; Indels 49; Gaps 5

Query	1	MILGLIASLVNGACPLMLPVLGEMSDNLI-----SGCLVQNTYSFPRL---	45
Db	49	malglalaahngltlpllnlvfygmtdstfqetrlplvtngselnstqvsdsleed	108
Qy	46	----TLVYGVIGVALLFETIOTISLMTITTAQOTKRIKQFHFVSLAODIGMFDSDDIE	101
Db	109	mamyayuytgyagvilyavlygslwclaaagrqhnlrkqkfhalmngelgwfdvndage	168
Qy	102	LNTFRMT-DIDKISDQIGDXTALLFQNMSPFSGILAVGLKGMKLTFLVLTSTPLIMASAA	160
Db	169	lntlrldvsklndqigdklmgfsgitlfsagfligfsgykltlvllavspilglsaa	228
Qy	161	ACSRWVLSLTSKELSNASXAGAAVEEVLSSIRTVIAFRAQEKELORSFLNITRYAMEFY	220
Db	229	mmakvlstfthkelqayakagavaeevlaaitrtvialafgqkkele-----	273
Qy	221	POWLISCVLAEVRYTONLKDADFGIKRTIKASVLSGANYFFPMNGVYGLAFWGTSLIN	280
Db	274	-----Tynklnleekrvqgkklatanlsfayllyvasayalaftwygtslvis	321
Qy	281	GEPTGTIGVLAVFPVSIIHSSYCIGAAPHFETFAIRGAAPFIPOVIDKKPSIDMFSTA	340
Db	322	ne-yslgvltvfislligtisiglhlapleefanaaraayelkildnepsidsfsk	379
Qy	341	GKPEISIEGTVFKNVSEVPSPRSIKIKGLNLRIKSEFVALVYLGNSGKSTVVQLQ	400
Db	380	ghkpsdimglnlefknyfrfypsersekllkglnlkkxsgqvalyvnsgcgsktvtqllq	439
Qy	401	RLYPDQDGIWDEMDIRALNVRHVRDHLGVVSOEPELVFEGTTSNNIKYGRDVTDEME	460
Db	440	rlldlpegsvsdqgdlrtlnvrylreilgyvsqpevlfatlaenlygfrentvnde	499
Qy	461	RAAREKNAVDLIMEFPNKNTLVGEGAGMSGQKORIAAALVAPNPKILLIDETSAL	520
Db	500	kavkeenaaydfimklphkfnlvergsaqls9gqkrtalatalvnpkllilldeatsai	559
Qy	521	DESKSAVOALEK 534	
Db	560	dteeseavvgaaldk 573	

RESULT 11  
 AAB60409  
 ID AAB60409 standard; Protein; 1275 AA.  
 AC AAB60409;  
 XX  
 DT 24-APR-2001 (first entry)  
 XX  
 DE Rat mdrlb2 multidrug resistance protein, SEQ ID NO:2.  
 XX

```

KW Rat; multidrug resistance; mdrlb2; multi-drug transporter family;
KM drug screening; pharmacokinetic analysis; oral absorption;
KM formulation design; bioavailability; transgenic animal; knockout animal;
KM inflammation; cardiovascular disease; central nervous system disorder;
KM cancer; autoimmune disorder; kidney disease.
XX
OS Rattus sp.
XX
PN US6169166-B1.
XX
PD 02-JAN-2001.
XX
PE 29-NOV-1999; 99US-0450105.
XX
PR 22-JUL-1998; 98US-0120513.
PA (SMIK ) SWITZERLINE BEECHEM CORP.
PI (SMIK ) SWITZERLINE BEECHAM PLC.
PT Chenery RJ, Ellens H, Brun KA, Yue L, Feld JA;
DR WPI: 2001-158280/16.
PS N-PSDB; AAF27498.
XX
PT New rat multidrug resistance protein (mdrlb2), useful for screening or
PT identifying compounds that are (ant)agonists/inhibitors of the
PT ratmdrlb2, as well as compounds with optimal development
PT characteristics -
XX
PS Claim 1; Column 17-24; 13pp; English.
XX
CC The invention relates to a novel rat multidrug resistance (mdr)
CC protein, mdrlb2 (AAB60409). Multi-specific drug transporter family
CC proteins are present in cell which have a barrier function, such as
CC intestinal epithelial cells, brain microvessel endothelial cells,
CC kidney epithelial cells, and liver hepatocytes, and are also expressed
CC by certain cancer cells. The rat mdrlb2 protein is useful for
CC screening or identifying compounds that are agonists or antagonists
CC of mdrlb2 activity. It may also be used to establish assays to
CC predict oral absorption and pharmacokinetics of drugs in humans, and
CC thus enhance the design of formulations through the identification of
CC compounds with optimal development characteristics (i.e., high oral
CC bioavailability, UDP (once a day) dosing, reduced drug interactions,
CC reduced variability, and reduced food effects), specifically to
CC avoid interactions with human mdr-1. Transgenic and knockout animals
CC created using DNA encoding the rat mdrlb2 may be used to gain an
CC insight into treating and preventing human diseases such as cancer,
CC inflammation, cardiovascular disease, central nervous system disorders,
CC autoimmune disorders and kidney disease. The present sequence represents
CC rat mdrlb2.
XX
SQ Sequence 1275 AA;
XX
Query Match 56.7%; Score 1552.5; DB 22; Length 1275;
Best Local Similarity 55.6%; Pred. No. 1.8e-142;
Matches 308; Conservative 92; Mismatches 105; Indels 49; Gaps 5
OY 1 MILGILASIVNCACPLAPLVGEKSDNLI-----SGCLVOTNTYSFRL--- 45
   I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
Db 49 malgilaaihlgtpllmvlfgymtstfgaetrlpsvtngseinstqtvsdsleed 108
   ||||| |I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
OY 46 ---TLVVVGICVALIFGYIOISLWIIITAQTRIRROPFHSVAIOIGFDSCDICE 101
   ||||| |I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
Db 109 mamayyytytgagvllvaaylqvswclaagqnlkltqrklthaamngelgwfdndage 168
OY 102 LATMT-DIDKIDSIGDKIALLFONMSSTFGSLAVGLVKGRKLTVLTSPSLIMASA 160
   ||||| :|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
Db 169 lntirltdvskindlgdklmgftsfitsgfllgfsygvkltlvlaavspjglssa 228
OY 161 ACSRWVISLSELSAYSKAGAAVEBVSSITPVIAFAROEKELORSFLINTRTAWTF 220
   ::::|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|I|
Db 229 maakiletfimtelalayagaavaevlaaitvfafggqqkele----- 273
```

Oy	221	PQMLLSCVLAFPRYQNQLKDAODPFCIKRTIASKVSLGANYFFPMNGTGYGLAEFWYGSLLIN	280
Oy	274	-----rtnknlleearvgyikkaitanlsiglayllvasyalatwygtslyvs	321
Oy	281	GEPGYTGIVTLAVFSPSVHSXICGAAPVHEFTFAIARGAAPHFOVIDKKPSINDFTA	340
Db	322	ne--ysisqgvltvfssilgtfslghlpnlaafanaagaafeiklndnepsidsfsck	379
Oy	341	GKPSEISGEYVEFKNVSEFNYPSPRSPIKILKGLNLRKSGEYVALVLGNSGKSTVVOLIO	400
Db	380	ghkpdsimgnlefknvyfnypsrsevkllkglnlkvksqglvalvgnsqcgkstctvqlq	439
Oy	401	RLYDDDFGFIAMDENDIRALNRHRDRDHIGVYSOEPVLFGTTISNNIKKGRDVMDEME	460
Db	440	Ilydilegevsidggdlrtlnvrylrrellgyvsqepvlattlaealirgrnvmdete	499
Oy	461	RAARAAANVADFEMPEPNKNTLVGERGAOMSGOKORIAIARALVRNPFIILIDEATSAL	520
Db	500	kavkeaneydfimrkphkfntlvgergaqlsggqkqrilatatalvrnpklllldetsal	559
Oy	521	DSEKSAVOALEK 534	
Db	560	dtesaavvgaaidk 573	
 RESULT 12			
ID	AAM48997	AAM48997 standard; Protein; 1280 AA.	
AC	AAM48997;		
DT	23-SEP-1998	(first entry)	
XX	Wild-type human P glycoprotein (PgP).		
KM	Human P glycoprotein; PgP; multi-drug resistance; cancer;		
KM	UIC monoclonal antibody; MAB; cytotoxic; transmembrane efflux pump.		
OS	Homo sapiens.		
PN	WO9821325-A1.		
PD	22-MAY-1998.		
PF	17-NOV-1997; 97WO-US21214.		
PR	15-NOV-1996; 96US-0752447.		
PA	(INGEN) INGENEX INC.		
PA	(UNIT ) UNIV ILLINOIS FOUN.D.		
PI	Mechetner E, Roninson IB;		
PI	WPI: 1998-297930/26.		
PT	N-PSDB; AAV32645.		
PT	Immunological reagent specific for P-glyco:protein - useful for		
XX	detecting multi-drug resistant cancer, isolating haematopoietic		
XX	cells and selective cell killing		
PS	Disclosure; Pages 38-42; 89pp; English.		
XX	The present sequence represents the wild-type human P glycoprotein		
CC	(PgP) protein. PgP is a transmembrane efflux pump protein involved		
CC	in multi-drug resistance of cancer cells. The invention provides		
CC	methods for developing and using immunological reagents specific for		
CC	certain mutant forms of PgP and wild-type PgP in a conformation		
CC	associated with substrate binding or in the presence of ATP depleting		
CC	agents. An example of the immunological reagent is the UIC2 monoclonal		
CC	antibody (mAb). mAb UIC2 specifically binds to PgP in a particular		
CC	biochemical conformation and is capable of inhibiting drug efflux from		
CC	PgP-expressing cells. The immunological reagents are claimed to be		

CC useful for detecting Pgp expression in mammalian cells, including low  
CC level expression, particularly in cancer cells to diagnose multi-drug  
CC resistance. The invention claims that these immunological reagents are  
CC more specific than known reagents for detecting Pgp and they also  
CC eliminate the need for costly and laborious screening of Pgp inhibitors  
CC by cytotoxicity or dye exclusion methods.

XX Sequence 1280 AA;

Query Match 56.7%; Score 1552.5; DB 19; Length 1280;  
Best Local Similarity 55.1%; Pred. No. 1.8e-142;  
Matches 305; Conservative 101; Mismatches 99; Indels 49; Gaps 6;

```

QY 1 MIIIGTASLVNGACLPMPVLVGLGMSD-----NLISGCLVQV--NTYSFP----- 43
DB 51 mvygtlaaiahgaglpimmlvfgemtdifanagnledlmsnltcrsdindgtffmleed 110
QY 44 -RLTLVYVGIGVAALFEGYQISLWITTAAROTKRIKOPFHSVLADIGWEDSCDIGE 101
DB 111 mtrayayysgigagvlyaaayqvsfwclaagrqhkirqgfihaimrgelgwfdvhdvge 170
QY 102 LNTRMV-DIDKISDGDGDKTALLFQNNSTFSGIAGVLVKGWKLTVLTSTPLIMASAA 160
DB 171 Intrltddvskinegldgkigmftgsmatftgfygftgwkltvlylalspylgisaa 230
QY 161 ACSRMVLSLSKELSAVSKAGAAVEVLSSIRVIYARAOKELOREFLNITRYAMFYF 220
DB 231 vwakllsfcdkellayakagavaeevlylaalrtvialfgqkkel----- 275
QY 221 PQWLSCVLNFRYTONLMDKADFGIKRTIASKVSLGAVYFEMNGTGLAFWGTSLILN 280
DB 276 -----rnhkneakrigrigkaltanlsigaalllyasvafwygttlvls 323
QY 281 GEGYITGTVLAFVFSYIHSSYCCGAAVPHFETFAIRGAFAHFQYIDKKRPSIDNSTA 340
DB 324 ge--ysyqglvlyfsvlylqafsvqgspisleaanaragaeyefkildkpsidsysks 381
QY 341 GYPRSEIEGVVEEFKFNPNPSRPSIKILKGLNLRKSGETVALVNGSGKSTVOLLQ 400
DB 382 ghpdnklgnlefnvhsfprskvklkglndkvsqgtvalvngsgqskstvtqlmg 441
QY 401 RLTPDDGFLTVDENDIRALNVRHRIHIGVSOEPLVFGTTSSNNIKYGRDVTDEME 460
DB 442 rlyptegmvsvdqdlrtlnvrlfreligvsgqepvlfatlaenlrygreavtmdete 501
QY 461 RAAREAAVYFIMFEPKFNPTVNGEKGAGMGGOKRIATARAIVRNPKITIIDEATSAL 520
DB 502 kavveanaydfimklpklfdtlvgergaqlsgqkqrlataraivrnpxkllldeatsal 561
QY 521 DSEKSAVQAALER 534
DB 562 dteseeavvqvaldk 575

```

RESULT 13  
AAW48999  
ID AAW48999 standard; Protein; 1280 AA.

XX AAW48999;  
XX 23-SEP-1998 (first entry)  
XX C-terminal single mutant human P glycoprotein (K1076M).  
XX Human P glycoprotein; Pgp; multi-drug resistance; cancer;  
XX UIC2 monoclonal antibody; mAb; cytotoxic; transmembrane efflux pump.  
XX Homo sapiens.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 1076

/note- "Changed from Lys in wild-type to Met in mutant"

MO9821325-A1.

22-MAY-1998.

17-NOV-1997; 97WO-US21214.

15-NOV-1996; 96US-0752447.

(INGE-) INGENEX INC.

(UNIT) UNIV ILLINOIS FOUND.

Mechetner E, Roninson IB;

WPI; 1998-297930/26.

Immunological reagent specific for P-glyco:protein - useful for detecting multi-drug resistant cancer; isolating hematopoietic cells and selective cell killing

Example 3; Page -: 89pp; English.

The present sequence represents the C-terminal single mutant human P glycoprotein (KM Pgp) protein. Pgp is a transmembrane efflux pump protein involved in multi-drug resistance of cancer cells. The invention provides methods for developing and using immunological reagents specific for certain mutant forms of Pgp and wild-type Pgp in a conformation associated with substrate binding or in the presence of ACP depleting agents. An example of the immunological reagent is the UIC2 monoclonal antibody (mAb). mAb UIC2 specifically binds to wild-type Pgp in a particular biochemical conformation and is capable of inhibiting drug efflux from Pgp-expressing cells. Mutants of Pgp were produced to investigate the relationship between Pgp function and UIC2 reactivity. Mutations were created at nucleotide-binding sites whereby lysine residues 433 and/or 1076 were substituted with methionine residues. Therefore the present mutant had the lysine residue at position 1076 substituted with a methionine residue whilst leaving the lysine residue at position 433 unsubstituted to create the KM Pgp mutant. The KM mutant showed similar UIC2 reactivity to the wild-type Pgp protein. The immunological reagents are claimed to be useful for detecting Pgp expression in mammalian cells, including low level expression, particularly in cancer cells to diagnose multi-drug resistance. The invention claims that these immunological reagents are more specific than known reagents for detecting Pgp and they also eliminate the need for costly and laborious screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.

Note: The present sequence was not given in the specification, but created using the information given in Example 3 and the wild-type Pgp protein sequence provided.

Sequence 1280 AA;

Query Match 56.7%; Score 1552.5; DB 19; Length 1280;  
Best Local Similarity 55.1%; Pred. No. 1.8e-142;  
Matches 305; Conservative 101; Mismatches 99; Indels 49; Gaps 6;

```

QY 1 MIIIGTASLVNGACLPMPVLVGLGMSD-----NLISGCLVQV--NTYSFP----- 43
DB 51 mvygtlaaiahgaglpimmlvfgemtdifanagnledlmsnltcrsdindgtffmleed 110
QY 44 -RLTLVYVGIGVAALFEGYQISLWITTAAROTKRIKOPFHSVLADIGWEDSCDIGE 101
DB 111 mtrayayysgigagvlyaaayqvsfwclaagrqhkirqgfihaimrgelgwfdvhdvge 170
QY 102 LNTRMV-DIDKISDGDGDKTALLFQNNSTFSGIAGVLVKGWKLTVLTSTPLIMASAA 160
DB 171 Intrltddvskinegldgkigmftgsmatftgfygftgwkltvlylalspylgisaa 230
QY 161 ACSRMVLSLSKELSAVSKAGAAVEVLSSIRVIYARAOKELOREFLNITRYAMFYF 220

```





PR 28-SEP-1999; 990S-0156510.  
XX  
PA (GENT-) GENTEST CORP.  
XX  
PI Stocker PJ, Steinel-crespi DT, Crespi CL, Reif TC, Patten CJ;  
XX  
DR MPI; 2001-235373/24.  
DR N-PSDB; AAD03489.  
XX  
PT New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT Pgp inhibitors -  
XX  
PS  
PS Example 2; Page 72-75; 11pp; English.  
XX  
CC The invention relates to dog P-glycoprotein (Pgp) also referred  
CC as multidrug transporter (MDR) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening Pgp inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by Pgp activity, by  
CC reducing or increasing Pgp activity in a cell. Pgp nucleic acids  
CC are used as oligonucleotide probes. Complements of Pgp nucleic  
CC acids are useful as antisense oligonucleotides, to induce a Pgp  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (Pgp). The  
CC Pgp enzyme functions as an efflux pump exporting small molecules  
CC across the cell membrane. This enzyme is a member of the ABC  
CC transporter family.  
CC  
CC  
SQ Sequence 1280 AA.

Query Match 56.7%; Score 1552.5; DB 22; Length 1280;  
Best Local Similarity 54.3%; Pred. No. 1.8e-142;  
Matches 302; Conservative 98; Mismatches 105; Indels 51; Gaps 5;  
QY 1 MIIIGILASLVNGACLPMLPLVGLGEMSDNLISGCL-----VQINTYSFPR--- 44  
DB 50 mlygtmaailhgaalpimmlvfgmtdsfanaqisrntkftpvlinesltntqfihle 109  
QY 45 -----LLLYVVGIGVALLIEGYIOISLIIITAAROTKRIRKOPFHSVLAODIGWFDSCDI 99  
DB 110 eemltvayyyysjagavvvaayqsvtwc:laagrqllkirkqfthaimrgeiygfdvhdv 169  
QY 100 GELNTRMT-DIDKISDGIKIALLFQNMSTFISGLAVGLVKGWKLTVLTSLTSPILMAS 158  
DB 170 gelntrltddvskinegldkvgmffgslatffglvfgfpgwklclvllaalspvlqsls 229  
QY 159 AAACSRWVISTSKELSAISKAGAAVEVLSSITVTIAFRAQEKELQRSFLNITRYAMF 218  
DB 230 aalwakllsftfdkellayakagavaevlaaitvtaifgqkkel----- 276  
QY 219 YFPQWLSCVLXEFVRYTONLKDAKDFGKRTIAKSVSLGAYFFPMNGTYGLAFWYGTSLI 278  
DB 277 -----rYnknlleaakrllgikaitanlsigaeilliyasyalatwygcslv 322  
QY 279 LNEPGYITIGTVLAVFFSVIHSSYCI GAAYPHFETFAIARGAAPHIIFOVIDKKPSIDNFS 338  
DB 323 lsse-yltqgvlcvffsvllgafslgqaapsleafanargaaeyelfkllnhkpsidsys 380  
QY 339 TAGYKPSISIEGTVEFKNVSFNYPSPRSIKILKGLNRKSGETVALVGLNCGSKSTVYQL 398  
DB 381 ksgnhpdknlgnefknhvfeyprkewkllkgnllkxqsgqvalvynsgcgkstvtql 440  
QY 399 LQRIYDPDGEIMDENDIRALNVRHRRDHIGVVSQEPVLEFGTISNNIKYGRDVTDEE 458  
DB 441 mqrlydpdcdgmvcldgqgdlclnvrhrlrelgvyvsqepvlfatctaenllyrgrenvnde 500  
QY 459 MERARERANAYDFIMEFPNKFNTLVGERGAOMSGOKORIAIRALVNRPKILLIDEATS 518

DB 501 Iekavkeanaydfimklpnkfdclivgerarlsgqgkqrialaralvrnpkllildeats 560  
QY 519 ALDSEKSAVQAALEK 534  
DB 561 aldtseavvqvaidk 576

Search completed: April 22, 2002, 14:34:22  
Job time: 99 sec

---

**THIS PAGE BLANK (USPTO)**



Db	1	YSFRLTLTYVYGIGVAALIFGYIDISLMIITAAQTRIRKQFHSVLADIGMFSDCDI	60
Qy	100	GELNTRMT-DIDKISDGIQDKIALLFQNMSTFSIGLAVGLYKGMKLLVTLSTSPILMAS	158
Db	61	GELNTRMTDIDKISDGIQDKIALLFQNMSTFSIGLAVGLYKGMKLLVTLSTSPILMAS	120
Qy	159	AAACSRWYISLTSKLELSAYSKAGAAVEVLSSIRTVIAFRAOEKELORSFLNTRYAMF	218
Db	121	AAACSRWYISLTSKLELSAYSKAGAAVEVLSSIRTVIAFRAOEKELO-----	167
Qy	219	YFPQMLSCVLAEVRYTONLKDAQDFGIKRTIASKVSLGAVYFPFMNGTYGLAEWYGTSLI	278
Db	168	-----RYTONLKDAQDFGIKRTIASKVSLGAVYFPFMNGTYGLAEWYGTSLI	213
Qy	279	LNIGEPGYIGTVLAVFVSFVHSYSCICGAAVHEFTPALARGAAHIFQVIDKRPISDNFS	338
Db	214	LNIGEPGYITGVYLVAFVSFVHSYSCICGAAVHEFTPALARGAAHIFQVIDKRPISDNFS	273
Qy	339	TAGYKPESIEGTEVEFKNVSFNYPSPRSIKILKGLNLRKISGETVALGLNGSGKSTYVOL	398
Db	274	TAGYKPESIEGTEVEFKNVSFNYPSPRSIKILKGLNLRKISGETVALGLNGSGKSTYVOL	333
Qy	399	LQRLYPDPPDGIIMDENDIRALNVRHTRDHIQVYSGQEPVLFGTTISNNITYGRDVTDEE	458
Db	334	LQRLYPDPPDGIIMDENDIRALNVRHTRDHIQVYSGQEPVLFGTTISNNITYGRDVTDEE	393
Qy	459	MERAAARENAADFEIMEFPNKNFTYVGEKGAQMSGGOKORIAIATALVRNKRILLIDATS	518
Db	394	MERAAARENAADFEIMEFPNKNFTYVGEKGAQMSGGOKORIAIATALVRNKRILLIDATS	453
Qy	519	ALDSEKSAVOALEK 534	
Db	454	ALDSEKSAVOALEK 469	
RESULT 2			
Q46605		PRELIMINARY: PRT: 1280 AA.	
AC	046605		
DC	01-JUN-1998 (TREMBLrel. 06, Created)		
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	MULTIDRUG RESISTANCE P-GLYCOPROTEIN.		
GN	MDR1.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_Taxid=9615;		
RN	[1]		
RP	TISSUE=COLON.		
RC	SEQUENCE FROM N.A.		
RA	Puel O., Lepage J.F., Alvinerie M., Gallier P., Pneau T.,		
RL	Submitted (Jun-1998) to the EMBL/GenBank/DBJ databases.		
CC	-I- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC		
CC	TRANSPORTERS).		
DR	EMBL: AF045016; AAC02113.1; -.		
DR	InterPro: IPR003593; AAA.		
DR	InterPro: IPR001140; ABC_transporter_tmem.		
DR	InterPro: IPR003439; ABC_transporter.		
DR	InterPro: IPR001687; ATP_GTP_A.		
DR	Pfam: PF00664; ABC_membrane; 2.		
DR	Pfam: PF00005; ABC_tran; 2.		
DR	SMART: SMO0382; AAA; 2.		
KW	PROSITE, PS00211; ABC_TRANSPORTER; 2.		
Q0	ATP-binding; transport.		
Q0	SEQUENCE 1280 AA; 141523 MW; 762DD5AF64C73306 CRC64;		

Query Match	56.7%;	Score 1552.5;	DB 6;	Length 1280;
Best Local Similarity	54.3%;	Pred. No. 6.7e-97;		
Matches 302;	Conservative 98;	Mismatches 105;	Indels 51;	Gaps 5;
OY	1	MILGTLASLVNACPLPMPVLYGEMSDNLIISCL-----VQNTNYSPFR---	44	

Db	50	MLVGMALIIHGAALPLMLVFGNNTDSFANAIGSRNKTFPVITINEITNNTQHFINHLE	109
Qy	45	-----LTLVYVIGVAAALIFGYIOISLWIIITPAARQTRIRKQFPHSVLAADIGFSDCI	99
Db	110	BEMTYVAVYYSIGAGVLVAAYIQVSEFMCILAGAQIILKIRQFPHALIMPEIGMFVDHV	169
Qy	100	GELNTRMT-DIDKISDGIGDKIALLPQMSFISGLAVGLYKGGKLLVLTLSPLMAS	156
Db	170	GELNTRLDDVSKINEGIDYVGMFPOSIAFTEFGTIVGTFPGKKLLVLTALISPYGLS	229
Qy	159	AAACGRMTVLSLSKELSVSKAGAAVEEVLSSIRTVIAFRAOELEORSFLNTRYAMF	218
Db	230	AAIMAKIISFTDKELLAYAGAAVEEVLAIITVIAFGQKKELE-----	276
Qy	219	YFPQWLSCVLXFNRYTTONLKADDFGIKRTIASKVSLGAYTFPMNCTYGLAFMYGSLI	278
Db	277	-----RYNNKLEEARIGIKNAITNISGAFAFLIYASVALAFMYGSLV	322
Qy	279	LNAGEGYITGVLAIVESVISHSYICAGAAVHEFTFAIARGAFHITQVIDKRPISDNES	338
Db	323	LSSE--YTIQGLVTFEFSVLIGAGSISGASPSIEAFANARCAAEIPIKIDNKPISDYS	380
Qy	339	TAGYPESTIEGVKERNSENYSPRSPSKILIKGLNLTIKSGEYALVGLNGSGKSTVQL	398
Db	361	KSGRHPDNIKGLERKNVHFSYPSRKEKYLILGLNLVQSQVALVAGNSGCGKSTVQL	440
Qy	399	LQRLTDPDDGFIMVENDIRALNRYHNDHIGVYSGEPEVLFGTTISNNIKYGRDVTDEE	458
Db	441	MQRLYDLPDGMACIDGQIRIKINRHLREITGVYSGEPVLFATTIAINIKRGRENYTME	500
Qy	459	MERAREANAYDFIMEFPNKNFTLVGEKGAOMSGGOKORIAIARALVNRPKILILDEATS	518
Db	501	IEKAVKENAAYDFIKPLPKNKDFTLVEERGARLSGGOKORIAIARALVNRPKILILDEATS	560
Qy	519	ALDSEKSAVQAALKE 534	
Db	561	ALDTESEAVALDK 576	
RESULT 3			
Q9JK64		PRELIMINARY; PRN; 1272 AA.	
AC	Q9JK64;		
DR	01-OCT-2000 (TREMblrel, 15, Created)		
DR	01-OCT-2000 (TREMblrel, 15, Last sequence update)		
DR	01-JUN-2001 (TREMblrel, 17, Last annotation update)		
DE	MULTIDRUG RESISTANCE PROTEIN 1A.		
GN	PGY1.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=WISTAR; TISSUE=LIVER;		
RA	Hoolveld G.J.F.E.J., Wlms J.W.J., Hagenbuch B., Jensen P.L.M.,		
RA	Meijer D.K.F., Muller M.;		
RT	*Cloning and functional characterization of the rat multidrug		
RT	resistance protein Mdr1a.;		
RL	Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC		
CC	TRANSPORTERS).		
DR	EMBL; AF257746; AAP69007.1; -		
DR	InterPro; IPR003593; AAA.		
DR	InterPro; IPR001140; ABC_transporter_tmem.		
DR	InterPro; IPR003439; ABC_transport.		
DR	InterPro; IPR001687; ATP_GTP_A.		
DR	Pfam; PF00064; ABC_membrane.2.		
DR	Pfam; PF00065; ABC_tran; 2.		
DR	SMART; SM00382; AAA.2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER.2.		
KW	ATP-binding; Transport.		

DR    zncfinger1: zincfinger, 1  
DR    Pfam: PF00664; ABC\_membrane, 2.  
DR    Pfam: PF00005; ABC\_tran, 2.  
DR    SMART: SM00382; AAA, 2.  
DR    PROSITE: PS00211; ABC\_TRANSPORTER, 2  
KW    ATP-binding; Transport.

SQ - SEQUENCE 1272 AA; 140326 MW; B5C1DF89DFC986C5 CRC64;

Query Match 56.5%; Score 1548; DB 11; Length 1272;

Best Local Similarity 54.8%; Pred. No. 1.3e-96;

Matches 301; Conservative 101; Mismatches 101; Indels 46; Gaps 5;

QY 1 MILGLASLVNGACLPPLMPLVGEKSDNLISGLVQTNMYSPFRL----- 46  
 DB 50 MLGLGLAIIHIGIALPLMALVFGDMTDSFAN--VGNNSMSSTYNAIDYAKLEDEMTYYA 107  
 QY 47 LYYVIGVAALIFGYIOISLWITTAAROTKRIKOPFHSVLADIGWEDSCDIGELNTERM 106  
 DB 108 YYYTIGAGVILVAYIOYSLMCLAGROHLKIQKPFHAIMROEVGFVDHVGELNTRFL 167  
 QY 107 T-DIDKIDGIGDKTALLFQNMSTFSGIANGVLVGKMLTVLTSTPLIMASAAACSM 165  
 DB 168 TDDVSKINEGIGDKIGMFQAMATFFGFIIGFTRGWKLTVLIALSPVLGSLAGIWAIXI 227  
 QY 166 VISLSKELSAYSKAGAAVEEVLSSIRTVIAFRAOEKELORSFLNITRYAMFFPPQML 225  
 DB 228 LSSFDKELQAYAKAGAAVEEVLALRTVIAFGQKKELE----- 267  
 QY 226 SCVLAFFRYTONLKDAKDFIKRTIASKVSIGAVYFFPMNGTYGLAFWYGTSLINGEPGY 285  
 DB 268 -----RYNNNLEBAKRLGIRKAITANISMGAFLLIYASVALAFWYGTSLIVISKE--Y 318  
 QY 266 TIGTVLAVFVSYSYICGAAPVPEFTFAIRGAAPHIPOVIDKPPSIDNSTGAKPE 345  
 DB 319 TIGQVLTFFSVLIGAFSGQASPNIEAFANARGAAYEVFSIIDDKPSTIDSPSKSGHKD 378  
 QY 346 SIEGVEEKNVSNFSPRSITKILKGLNLRKSGEVALVGLNGSGKSTVOLLQRLYDP 405  
 DB 379 NIQGLLEERKNHFSTPSKRDVOILKGLNLRKVSQGTVALVNGSGKSTVOLLQRLYDP 438  
 QY 406 DDGEIMVDENDIRALNVHRDHIGVNSQEPVLEGTTSNNIKYGRDVTDEMERARAE 465  
 DB 439 IEGEVSIDGODIRTNVRLREILIGVSEPELFTTIAENIRYGRNVTMEIKAVME 498  
 QY 466 ANAYVFIMEFPNKENTLVGEKAGQMSGGOKORIAIARALVRNPKILLIDEATSALDESSEK 525  
 DB 499 ANAYVFIMKLPKFPDLVGERGAQLSGGOKORIAIARALVRNPKILLIDEATSALDESSE 558  
 QY 526 SAVOALAEK 534  
 DB 559 AVOALAEK 567

RESULT 4  
 ID 09TSU2 PRELIMINARY; PRT; 1163 AA.  
 AC 09TSU2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE MULTI-DRUG RESISTANCE RELATED MRNA, PARTIAL CDS (FRAGMENT).  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Okada Y., Nakamura N., Matsusiro H., Kato H., Setoguchi A., Yazawa M.,  
 RA Okuda M., Wataai T., Hasegawa A., Tsujimoto H.;  
 RT "Molecular analysis of multidrug resistance in lymphoma cells in the  
 cat.";  
 RL Am. J. Vet. Res. 0:0-0(1999).  
 CC -I- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 CC EMBL; AB029153; BAAB7071.1; -.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmam.  
 DR InterPro: IPR003439; ABC\_transporter.

DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF000664; ABC\_membrane\_2.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW ATP-binding; Transport.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 1163 AA; 128510 MW; 1B5B413776A93A26 CRC64;

Query Match 56.2%; Score 1540; DB 6; Length 1163;

Best Local Similarity 55.6%; Pred. No. 4.1e-96;

Matches 303; Conservative 93; Mismatches 99; Indels 50; Gaps 5;

QY 11 NCACPLMPLVIGEMSDNL-----ISGLVQTNMYSPFRL-----TLYYV 50  
 DB 1 HGAALPLMALVFGDMTDSFANISRLNLTNLITGESIYNDSTFIRLREEMTVAYYVS 60  
 QY 51 GIGVAAALIFGYIOISLWITTAAROTKRIKOPFHSVLADIGWEDSCDIGELNTRMT-DI 109  
 DB 61 GIGAGVLAALVYQVSWCLAGROHLKIQKPFHAIMROEVGFVDHVGELNTRFLTDVY 120  
 QY 110 DRISDIGDKTALLFQNMSTFSGIANGVLVGKMLTVLTSTPLIMASAAACSRMYISL 169  
 DB 121 SKINEGIGDKIGMFQSMATFFIGFTRGWKLTVLIALSPVLGSLAIAWAKILTSF 180  
 QY 170 TSKELSAYSKAGAAVEEVLSSIRTVIAFRAOEKELORSFLNITRYAMFFPPQMLSCVL 229  
 DB 181 TKEELAVAKAGAAVEEVLALRTVIAFGQKKELE----- 216  
 QY 230 XAVRTQNLKDAKDFIKRTIASKVSIGAVYFFPMNGTYGLAFWYGTSLINGEPGYTGT 289  
 DB 217 ---RNNKLEBAKRLGIRKAITANISIGAFLLIYASVALAFWYGTSLIVISHE--YSIQ 271  
 QY 290 VLAVFVSYSYICGAAPVPEFTFAIRGAAPHIPOVIDKPPSIDNSTAGAKPEISIG 349  
 DB 272 VLTFFSVLIGAFSGQASPNIEAFANARGAAYEIKFIIDNKPSTIDSKNGKPNING 331  
 QY 350 TVEFKNVSNFSPRSITKILKGLNLRKSGEVALVGLNGSGKSTVOLLQRLYDDGEF 409  
 DB 332 NIEFKNVHSPSKREVKILKGLNLRKVSQGTVALVNGSGKSTVOLLQRLYDDTDM 391  
 QY 410 IMVDENDIRALNVHRDHIGVNSQEPVLEGTTSNNIKYGRDVTDEMERARANAY 469  
 DB 392 VSIDGODIRTNVRLREILIGVSEPELFTTIAENIRYGRNVTMEIKAVKANAY 451  
 QY 470 DEIMEFPNKENTLVGEKAGQMSGGOKORIAIARALVRNPKILLIDEATSALDESSESAVO 529  
 DB 452 DEIMKLPKFPDLVGERGAQLSGGOKORIAIARALVRNPKILLIDEATSALDESSEAVVO 511  
 QY 530 AALEK 534  
 DB 512 VALDK 516

RESULT 5  
 ID 002793 PRELIMINARY; PRT; 1285 AA.  
 AC 002793;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN-1.  
 GN MDR1.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Longley M., Crawford A.M.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
CC TRANSPORTERS).  
DR EMBL: U78609; AAB58489.1; -.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR APB-binding; Transport.  
KW SEQUENCE 1285 AA; 142020 MW; 90153617C44856F CRC64;  
SQ  
Query Match 55.7%; Score 1525; DB 6; Length 1285;  
Best Local Similarity 53.6%; Pred. No. 5e-95;  
Matches 300; Conservative 100; Mismatches 100; Indels 60; Gaps 6;  
OY 1 MILGILASLVNGACLPMLPLVLGEMSDNLIS-----GCLVQTN----- 38  
DB 55 MVLGTLAIIHAGLPLMLVLFGDMTDSFAGAGNLGNTLISNTSTIDRTGKALEK 114  
OY 39 ---TYSEFRLLTYVYGVAALIFGYIOISLMTITTAARQTRIRKQPFHSYLAODIGMFD 95  
DB 115 EMITYAY-----YSSGIGAVLIAIYQVSWCLAGRQVRIRKQFHALMQEIGMFD 169  
OY 96 SCDIGELNTRMT-DIDKISDGIGKIALLFQNMSTFSIGLAVGLVKKLLVLTSTPL 154  
DB 170 VHDGAKLNTRLTLDVSKINEGIGKIGMFPQAMATFFFTGFTIGFTTGLTLVLIATIPV 229  
OY 155 IMASAACSRSVILSTKELSAVSAVAEEVLSSIRTVIAFRAOEKELQRSPLNTR 214  
DB 230 LGLSAIATKALISFTDKKLALYAKAGAAVEEVLATITVIAFGQKKELE----- 280  
OY 215 YAMFFPQWLLSCVLFVRYTQNLKDAKDFGIKRTIASKVSLGAVYFPNMGTYGLAFWYG 274  
DB 281 -----RYKNLEEKARIGIKKAITANISGAALFLIYASALAFWYG 322  
OY 275 TSLINGPGTIGTVLAVFVSIIHSSYCIGAAPHFETFAIARGAAHIFQVIDKRSI 334  
DB 323 TSLVLSRE--YSIQOVLFVFSVLIGAFSIGQASPNIEAFANARAAVEVERKIIDNKPSI 380  
OY 335 DNFSAGAKPESIGVEFKNVSEFNPSPSIKILKGLNLRIKSGETALVGLNGSGST 394  
DB 381 DSYNTGKPKPNIGNENLEFRVNHRYHPSRNEVKILKGLNLVSGQYVALVNGSGCKST 440  
OY 395 VVOLLQRLYDDDFIWDENDIRALNRYRHDHIGVVSQEPVLEGTITSNIKYGRDV 454  
DB 441 TVQLMQRLYDTEGVNSIDGQDIFIRNRYLREIIGVVSQEPVLEFATITAEIRYGRDV 500  
OY 455 TDEMERARARANAVDFMEFPNKENTLVGKGAQMSGGQKORIAALAVRNPKILLD 514  
DB 501 TMDIQAANKKANNYDFIMKLPNKFDVLVGERGALSGGQKORIAALAVRNPKILLD 560  
OY 515 EATSALESKSAVOALEK 534  
DB 561 EATSALESSEAVVOALDK 580  
RESULT 6  
OY1586 PRELIMINARY; PRT; 1287 AA.  
AC OY1586;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN.  
GN XEMDR.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;

OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8335;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-95322451; PubMed-7599185;  
RA Castillo G., Shen H.J., Horwitz S.B.;  
RT "A homologue of the mammalian multidrug resistance gene (mdr) is  
RT functionally expressed in the intestine of Xenopus laevis.";  
RL Biochim. Biophys. Acta 1262:113-123(1995).  
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
CC TRANSPORTERS).  
DR EMBL: U17608; AAA75000.1; -.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
DR APB-binding; Transport.  
KW SEQUENCE 1287 AA; 141504 MW; 0695073C5771415 CRC64;  
SQ  
Query Match 55.6%; Score 1523; DB 13; Length 1287;  
Best Local Similarity 54.6%; Pred. No. 6.9e-95;  
Matches 300; Conservative 91; Mismatches 114; Indels 44; Gaps 4;  
OY 1 MILGILASLVNGACLPMLPLVLGEMSDNLISGCLVQTNFYSEFR-----T 46  
DB 66 MLFGTIALSAGAAALPLMLVLFGEMTDSFVNVGQVDGNTFWESMINASRELOGMTTYA 125  
OY 47 LYYVIGVAALIFGYIOISLMTITTAARQTRIRKQPFHSYLAODIGMFDSCDIGLNTRM 106  
DB 66 MLFGTIALSAGAAALPLMLVLFGEMTDSFVNVGQVDGNTFWESMINASRELOGMTTYA 125  
OY 126 YYSGLGVMCAITQISEFTLSAGRQIKRSNFFFAVLRQETGWDINDAGLNTRL 185  
DB 107 T-DIDKISDGIGKIALLFQNMSTFSIGLAVGLVKGRLTVLTSTPLIMASAACSRL 165  
OY 186 TDDVSKINEGIGDKTAMLLQSLTTLVTFIIGTIGKMLTWVGMALSPIMGSAIAIMKY 245  
DB 246 LSAFNKELKAKYAKAGAAVEEVLSSIRTVFAFGQKKEIH----- 285  
OY 226 SCVLKFFVRYTQNLKDAKDFGIKRTIASKVSLGAVYFPNMGTYGLAFWYGTSLINGEPY 285  
DB 286 -----RIEKNLEDAKIGIKKAITANVSGFAFLMTIYAASLAFWYGTILLIDG--GY 336  
OY 286 TIGTVLAVFVSIIHSSYCIGAAPHFETFAIARGAAHIFQVIDKRSIDNFSAGYKPE 345  
DB 337 TIGSVLVFPAVILIGAFVAGQTSPIIEAFANARGAATITFNIDQPKIDSRSKGLKPD 396  
OY 346 SIEGVEKKNVSENPSPSIKILKGLNLRIKSGETVALVGLNGSKSTVQOLLQRLDP 405  
DB 397 KIKGDIKFNKVIIFYPSRKDIOVLKGLNLINPSGTVLAVSSGCKSTVQOLLQRFDP 456  
OY 406 DDGFTMVENDIRALNRYRHDHIGVVSQEPVLEGTITSNIKYGRDVQTEEMERARE 465  
DB 457 EDGVTTLTGQDIRSLNRYLREIIGVVSQEPVLEFPTTADINIRYREVTAEIIRAKE 516  
OY 466 ANAYDFIMEFPNKENTLVGKGAQMSGGQKORIAALAVRNPKILLIDEATSALESSEK 525  
DB 517 ANAYDFIMKLPDKLETVLVGERGTOLSGGQKORIAALAVRNPKILLIDEATSALEPTESE 576  
OY 526 SAVQALEK 534  
DB 577 AVVQSALDK 585  
RESULT 7  
OY3437 PRELIMINARY; PRT; 1287 AA.  
ID OY3437  
AC OY3437;

DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE ABC TRANSPORTER PROTEIN.  
 GN CMDRI.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=INTESTINE.  
 RA Edelman H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C.,  
 RA Kane S.E., Kuchler K.;  
 RT "Cmdri, a chicken p-glycoprotein, confers multidrug resistance and  
 RT mediates estrogen transport."  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 DR EMBL: AJ009799; CAA0835.1; -.  
 DR HSP; P13569; INBD.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR InterPro: IPR00504; RRM.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR PROSITE: PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1287 AA; 141785 MW; 3CPDBA8D657111F CRC64;

Query Match 54.2%; Score 1485; DB 13; Length 1287;  
 Best Local Similarity 53.0%; Pred. No. 2.6e-92;  
 Matches 292; Conservative 95; Mismatches 118; Indels 46; Gaps 5;  
 QY 1 MILGLASLVGACLPFLPVLGEMNDLI-----SGCLVQNTYS-----FFR 44  
 DB 61 MIFSLAIANGTSLPIMILFGDMTDFVSGMNTINGSSGLNSADVKNKLEEMTR 120  
 QY 45 LTLVYGVGVALIFGYIOISLMTITTAARQTRIRKOPFHSVLAODIGFDSCDIGELNT 104  
 DB 121 YAYYSALAAVLAAYIQTSTFLAAGROYKIRKREFFHAIMROELQMFVNDGELNT 180  
 QY 105 RM-TRDIDKISDIDGDKIALFQNMSTFSGLVGLVKGWKLTLVLTSLPLMASAACS 163  
 DB 181 RLIDVSKINIGDKIGFLIOSETFLTGFTIGVIRGSKLTLVLAVSPVLGSLAALMA 240  
 QY 164 RMVLSLTKELSAVSKAGAAVEVLSSIRTVIAFRAOKELOKRSFLNITRYAFWYFQW 223  
 DB 241 KILAFETKDEQAALAKAAGAAVEVLASAVRYIAFGQKEIK----- 282  
 QY 224 LLSCLVAVRYTQNLKDAKDFGIRKTASKVSLGAVYFPMNGTYGLAFWYGTSLILNGEP 283  
 DB 283 -----RYHKNLEDAKRIKIRKAITSNISMGAAFLLIYASVALAWYGTTLILANE 332  
 QY 284 GYTIGTVLAVFEFSYHSSYCGAAPHFETPAIARGAAPHFOVIDKRPISDNSTAYK 343  
 DB 333 -YSIGNVLTVEFSVLGAFSIGOTAPSTIEAFNAGAAVAIFNIIDNPEIDSYSDAGHK 391  
 QY 344 PESTIGTEFFNVSPSPRSIKILKGLNIRIKSGETVALVGLNGSKSTVVOLORLY 403  
 DB 392 PDHIGKNLEFNVSPSPRSIRPDELTKLNLKLVNGGOTVALVGGCGCKSTTVOLQIFY 451  
 QY 404 DPDDGFLVNDIRALVNRHROHIGVSOEPLVFGTTISNNIKYGGDVTDEEMRAA 463  
 DB 452 DKREGTITIDODKSLNVRRLREITGVNOBPLVFATITENIRYGRDVTMEIERAT 511  
 QY 464 REANAVDPTIMEFPKFTLVGEKAGOMSGOKORIAIRALVRNKKILIDENATSLDSE 523

DB 512 KEANAVDPTIMKPKFEYVGEKAGOMSGOKORIAIRALVRNKKILIDENATSLDTE 571  
 QY 524 SKSAVQAALEK 534  
 DB 572 SESVQAALEK 582  
 RESULT 8  
 ID 060502 PRELIMINARY; PRT; 1169 AA.  
 AC 060502;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE P-GLYCOPROTEIN.  
 GN GGP-1.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC Cricetus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG.  
 RX MEDLINE=91154265; PubMed=1671863;  
 RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;  
 RT "Full length and alternatively spliced pgpl transcripts in multidrug-  
 RT resistant Chinese hamster lung cells."  
 RL J. Biol. Chem. 266:4545-4555(1991).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 DR EMBL: M59254; AAA37005.1; -.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1169 AA; 128938 MW; 72E25B7CE29DC185 CRC64;

Query Match 53.8%; Score 1473.5; DB 11; Length 1169;  
 Best Local Similarity 57.1%; Pred. No. 1.4e-91;  
 Matches 284; Conservative 90; Mismatches 88; Indels 35; Gaps 4;  
 QY 39 TYSFRLTYVGVGVALIFGYIOISLMTITTAARQTRIRKOPFHSVLAODIGFDSCD 98  
 DB 3 TYAV-----YYTGIGAGVLAIVYIOVSFWCLAAGQIRKIRKREFFHAIMROELQMFVND 57  
 QY 99 IGEIUTRMT-DIDKISDIDGDKIALFQNMSTFSGLVGLVKGWKLTLVLTSLPLMA 157  
 DB 58 VGLNTRFLTDVSKINIGDKIGFLIOSETFLTGFTIGVIRGSKLTLVLAVSPVLGSL 117  
 QY 158 SAACSRVVISLTKELSAVSKAGAAVEVLSSIRTVIAFRAOKELOKRSFLNITRYAW 217  
 DB 118 SAGIWAKTLSFTDKELOAYAKAGAAVEVLAIKRTIAFAGQKELE----- 165  
 QY 218 FYFQWLLSCLVAVRYTQNLKDAKDFGIRKTASKVSLGAVYFPMNGTYGLAFWYGTSL 277  
 DB 166 -----RYNNNLEEAARKLIRKAITSNISMGAAFLLIYASVALAWYGTSL 210  
 QY 278 ILNGPGYTVLAVFEFSYHSSYCGAAPHFETPAIARGAAPHFOVIDKRPISDNSTAYK 337  
 DB 211 VLSKE--YSIGVLTVEFSVLGAFSIGOTAPSTIEAFNAGAAVAIFNIIDNPEIDSYSD 268  
 QY 338 STAGIKPESIEGTEFFNVSPSPRSIKILKGLNIRIKSGETVALVGLNGSKSTVVO 397  
 DB 269 SKNGYKPPNINIGNLEFNVSPSPRSIRPDELTKLNLKLVNGGOTVALVGGCGCKSTTVQ 328



[illegible]

```

Db          335  FCYALAFKPSYSTLVLD--EEETTPETVIOGIFLCVILAAAMNIGHMSSCSEIETSTGCSATNI 3933
OY          325  FQVIDAKPSIDNESTAGYKPSIEGIVFEKFNVSNNYSRPSIKLGNLRITSGEFTVAL 364
Db          394  FQITIDRQFVIDCMGSDGKYLDRIGETIEFHWTFPHYSRDPVKILDMLSVNIKPGETTAL 4533
OY          385  VGLNMSGKSTVYVOLLQRIYLPDDGFIWVDENDIRALNVRHYRHOIGVVSQEPVLFGTTIS 444
Db          454  VGSSGAGCKSTVALQILQRIYDDCEGMVWLDRSLIRMLRQDGIIVBEQVLPFTTIA 513
OY          445  NNIKGGRDQVDVDEMERARAPANAYDFIMEPHNKFTLVGEKGQMGSGCGKORITAIRAL 504
Db          514  ENIRGGRDADIMEDIQAAKMANNNYNTIMLPQGFDTLVGEGGQMGSGGQKQVATIRAL 573
OY          505  VRNPKILLIDEATSAIDSEKSAVOALEK 534
Db          574  IRNPKILLIDMATSAIDSESEARVOEALNK 603

RESULT      10
O45721      PRELIMINARY;          PRT;      1294 AA.
AC          045721      O62101;
DT          01-JUN-1998 (TREMBLrel, 06, Created)
DT          01-NOV-1998 (TREMBLrel, 08, Last sequence update)
DT          01-JUN-2001 (TREMBLrel, 17, Last annotation update)
DE          C47A10.1 PROTEIN.
GN          C47A10.1.
OC          Caenorhabditis elegans.
OC          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC          Rhabditidae; Peloiderinae; Caenorhabditis.
CX          NCBI_TaxID=6239;
LN          [1]
RP          SEQUENCE FROM N.A.
RA          Basham V.;
RL          Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN          [12]
RP          SEQUENCE FROM N.A.
RX          MEDLINE=94150718; PubMed=7906398;
RA          Wilson R., Ainscough R., Anderson K., Baynes C., Betks M.,
RA          Bonfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A.,
RA          Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA          Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA          Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA          Lightling J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA          Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA          Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA          Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA          Watson A., Weinstock L., Wilkinson-Sprat J., Wohlhahn P.,
RT          2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT          elegans.
RT          Nature 368:32-38(1994).
RL          [13]
RP          SEQUENCE FROM N.A.
RA          Basham V.;
RL          Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC          -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC          TRANSPORTERS).
CC          EMBL; 293782; CAB07855.1; -.
DR          EMBL; 281484; CAB07855.1; JOINED.
DR          EMBL; 281484; CAB03973.1; -.
DR          EMBL; 293782; CAB03973.1; JOINED.
DR          InterPro; IPR003593; AAA.
DR          InterPro; IPR001140; ABC_transporter_tmam.
DR          InterPro; IPR003439; ABC_transportr.
DR          InterPro; IPR001687; ATP_GTP_A.
DR          InterPro; IPR000130; Zn_Mtpeptide.
DR          Pfam; PF00664; ABC_membrane_2.
DR          Pfam; PF00065; ABC_tran; 2.
DR          SMART; SMO0382; AAA; 2.
DR          PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR          PROSITE; PS00142; ZINC_PROTEASE_UNKNOWN_1.

```

KW APP-binding; Transporth.  
 SQ SEQUENCE 1294 AA: 142490 MW: DBED2602PA484DF CRC64:  
 Query Match 41.4%; Score 1134.5; DB 5; Length 1294;  
 Best Local Similarity 42.8%; Pred. No. 1.8e-68;  
 Matches 235; Conservative 95; Mismatches 174; Indels 45; Gaps 5;

1 MILGIASLVGACPLMPVLYGEMSDNIS--GCLVOTN-----TYSFRL 45  
 46 LAVGIIVSCANGVGLPMSITIMGNVSNQFVTLGITFLDPNSTASEKAARAEFEVIQN 105  
 46 TLYVVGISGVAALIFGYIOISLMTTAAKOTRIKROFESHVLAODIGWFSDCIGELNTR 105  
 106 CLKVYVLCGGIFPAAGFLDASCFWICELSNRFRQFHSVMROELIAYDKNTSGTSLSNK 165  
 106 MTD-IDKISDGIKIALLFQNMSTFSGIAGLVKGMKTLVTLSPLIMASAAACSR 164  
 166 LFDNLERYREGTGDVGLAFQMAQOIFGFAVAFYDWLTLIMMSLSPFMWICGLFLAK 225  
 165 MVISLTSKELSAISKAGAVAEVLSITRTVIAFRAQEKELORSFLNITRYAMFFPWL 224  
 226 LLAATAAKREAKQYAVAGIAEEVLTSLRTVIAFNGQYECK----- 266  
 225 LSCVLXFRYTONKDAKDFGIKRTIASKVSLGAVYFEMNGTGYLAEWGTSLNGEPG 284  
 267 -----RYDALLEHGKKTGIKKSFLIGAGLAEFFVITVASTCLAEVGTNFVSGR-- 316  
 285 YTGTVLAVFESVISHSSYICGAAPHEFPAIARGAAPHIHOVIDKRPIDNFSTAGYKP 344  
 317 LESTGLVLFESVMMGSMALGOAGQOFATICTALGAASLVEYIDRPEIDAYSTEGQTP 376  
 345 ESIGTYEKNVSNYSPRSIKILKGLNLKSGEYVALVGLNGSGSTVYOLLQRLYD 404  
 377 SKISGRISVNRKVEETPTPRADVKILKGVSLDAPQGTVALVGGSCGKSTIIQLQRFYN 436  
 405 PDDEFIMVDEMDIALNVRHRYRDIHGVVSOEPVLEGTISNNIKYGRDVTDEEMERAR 464  
 437 PDAQIILDDLPIDENIKRIKQLVGVVSGQEPNLFNTSEONIRKGRSDVSEDIARLAK 496  
 465 EANAADFIEMEPNFKNTLVGKGAQMSGQKORIAIARALVRNPKILLDEATSLADES 524  
 497 EANAADFIEMEPNFKNTLVGKGAQMSGQKORIAIARALVRNPKILLDEATSLADES 556  
 525 KSAVOALE 533  
 557 ESIVOSALE 565

RESULT 11  
 001495 PRELIMINARY; PRT; 1289 AA.  
 AC 001495;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE SIMILARITY TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY.  
 GN C34G6.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peioderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RA [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; Pubmed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes T., Cooper J., Coulson A.,  
 Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,  
 Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 Gardner A., Green P., Hawkins T., Hillier L., Jler M., Johnston L.,  
 Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,  
 Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,

RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Woldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RL Langston Z., Rohlfing T.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RC STRAIN-BRISTOL N2;  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS).  
 DR EMBL: U97407; AAB52482.1; .  
 DR HSSP; P13569; INBD  
 DR InterPro; IPR003593; AAA.  
 DR InterPro; IPR001140; ABC\_transporter\_tmem.  
 DR InterPro; IPR003439; ABC\_transportr.  
 DR InterPro; IPR001687; ATP\_GTP\_A.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW APP-binding; Transporth.  
 SQ SEQUENCE 1289 AA: 143065 MW: AE2F6857ACFC6DDC CRC64:

Query Match 38.7%; Score 1060; DB 5; Length 1289;  
 Best Local Similarity 40.4%; Pred. No. 2.1e-63;  
 Matches 224; Conservative 102; Mismatches 180; Indels 48; Gaps 6;

1 MILGIASLVGACPLMPVLYGEMSDNIS--GCLVOTN-----TYSF-- 42  
 50 LITGVAAVHAGCPPLAIVLGMTVFLRAQNSDFVAGVDNVPBGLVPLISDEFNSE 109  
 43 -FRLLVYVIGVVAALIFGYIOISLMTTAAKOTRIKROFESHVLAODIGWFSDCIG 101  
 110 VKYCIYVYLVGLMFLFSYQIACFESYAEELVHKLNQNTLKALRLQIQWFDKQOQGN 169  
 102 LNTNMT-DIKISDGIKIALLFQNMSTFSGIAGLVKGMKTLVTLSPLIMASNA 160  
 170 LTAFLTDLERYREGIGDKFALLVQMFAPLAGYGVGFYSMTLVMMGFAPLIVLSGA 229  
 161 ACSRWVLSLTSKELSAISKAGAVAEVLSITRTVIAFRAQEKELORSFLNITRYAMFFY 220  
 230 KMSKMATRTREOETVAVAGIAEETSSIRTVHSLNGHRELDK-----FY- 277  
 221 POMLSLVXLFVRYTONKDAKDFGIKRTIASKVSLGAVYFEMNGTGYLAEWGTSLIN 280  
 278 -----NALEVGQRTGIYKCYMGIGTGVSNCLMSSYALAFYSGTSLIIN 322  
 281 GEPGTYIGTVLAVFESVISHSSYICGAAPHEFPAIARGAAPHIHOVIDKRPIDNFST 340  
 323 -DPTDRGLIFTFEFAVLSGSLGALPHLASFGTAGAATVLRVINSHPKIDPYSL 381  
 341 GYKPESTIGYEFKNVSNYSPRSIKILKGLNLKSGEYVALVGLNGSGSTVYOLLQ 400  
 382 GILVNMKGDSIFKDVHFRYPSRKDIHVLKISTELKAGDKTALVSGSGCGSTIVNLLQ 441  
 401 RLYPDDGFIMVDEMDIALNVRHRYRDIHGVVSOEPVLEGTISNNIKYGRDVTDEEME 460  
 442 RPYDPTKRVILIDVDLAEVNVHSLRQDIGISQEPVLEFGTYINIMKGNHATHDQV 501  
 461 RAARANAADFIEMEPNFKNTLVGKGAQMSGQKORIAIARALVRNPKILLDEATSL 520  
 502 EACKMANANDFKRLPDYDGVFRVGEKGVOLSGQKORIAIARALVRNPKILLDEATSL 561  
 521 DSEKSAVOALEK 534



DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran\_2.  
 DR SMART: SM00382; AAA\_2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1275 AA; 142203 MW; 7439D3836B75787C CRC64;

Query Match 37.5%; Score 1028; DB 5; Length 1275;  
 Best Local Similarity 40.5%; Pred. No. 3.1e-61;  
 Matches 223; Conservative 94; Mismatches 186; Indels 48; Gaps 6;

QY 4 GILASLVNGACLPMPVLVLEGM-----SDNLISGL-----VOTNTYSF-----R 44  
 DB 54 GTLAPCHGAGFSLVIGMTVPLRAQNSEFVLGVSDPGLPALTRKEFDTLVRR 113  
 QY 45 LTLTYVGVAALFYGIQISLWITTAARQTKRIRKOFHSHVLAQDIGWEDSCDIGELNT 104  
 DB 114 YCLYVLGIFGFMAFATSYIOICWETFEERITHKRLIKLITLKAIRKQISWPFIDQGTGNLTA 173  
 QY 105 RMT-DIKISGIDGDKTALLFQNMSTFSGIAGVLGKWKLTLYLTSTPLMASAACS 163  
 DB 174 RLTDLEVRERGLDGLSLFQVSAFVAGFCVGFANSMSTLVMMVAVAPVIVISANMS 233  
 QY 164 RMVLSLTKELSAVSKAGAAVEVLSIRTVIAFRAOEKELQSRFLNITRYAMFYEPQW 223  
 DB 234 KIVATPQVEGETYAVAGALAEETFSIRIVSHICGHRKL----- 274  
 QY 224 LLSCLVLRVTRQNLKDAKDGIRKIRTSKVSLSGAVYFPMNGTGLAFWYGTSLINCEP 283  
 DB 275 -----TREALELGRQGRQGLVKKYFVGVGEGQCTVSYALAWYSVLIN-DP 325  
 QY 284 GYTITVLAVFESYIHSSYICGAAPHEPFIARGAFFHFOVDKPKSIDNFSTAGYK 343  
 DB 326 ALDGRITTVFAVMSGSAALGTCLPLHNTISIRAGVRSVLSYNSPKIDPSLSDIV 385  
 QY 344 PESIEGVEFNKNSFNPSRPSIKLKLNLRIKSGETVALVGLSGKSTVYVOLLRLY 403  
 DB 386 LNNKRSIRFKNVHPSYRSRTLDLKGVSLOVSAGKIALVSSGCKSINVMLLRFY 445  
 QY 404 DPDDGFIVDENDIRALNVRHRYRDHIGVSOEPLVFGTTISNNIKYGGDDVTDDEMEAA 463  
 DB 446 DPTGRKVIDIDIDVCDLWVKLRQIGVSOEPLVFGTTLEENIKMGYEOATMEVQDAC 505  
 QY 464 REANAYDIFMEFNKNTLVGEKGAOMSGOKORAIARALVRNPKLILIDEATSALDSE 523  
 DB 506 RVANNAADTKRLPEGYGRVGERGVOLSGGOKORAIARAIINPKRILLDEATSALDTE 565  
 QY 524 SKSAVOALEK 534  
 DB 566 AESTVOALEK 576  
 RESULT 14  
 Q9LGLX1 PRELIMINARY; PRT; 1285 AA.  
 AC Q9LGLX1;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE SIMILAR TO ARABIDOPSIS THALIANA CHROMOSOME 2.  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Ehrhartoideae; Oryzaceae; Oryza.  
 OK NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone: p0706B05. ";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 DR EMBL: AP002482; BAA96612.1; -  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran\_2.  
 DR SMART: SM00382; AAA\_2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1285 AA; 138049 MW; 2D45FEC7E956A1A0 CRC64;

Query Match 36.9%; Score 1009.5; DB 10; Length 1285;  
 Best Local Similarity 38.8%; Pred. No. 5.6e-60;  
 Matches 208; Conservative 112; Mismatches 185; Indels 31; Gaps 4;

QY 1 MILGILASLVNGACLPMPVLVLEGMDSNLISGLVOTNTYSFRLTYVVGVAALIFG 60  
 DB 71 MALGTIGAVANGALPFWTVLFGNLIDAFGAGMGIHGVNVSWSLEFIYLAISAVAS 130  
 QY 61 YIQISLWITTAARQTKRIRKOFHSHVLAQDIGWEDS-CDIGELNTRMT-DIDKISDIGD 118  
 DB 131 FVOVTCWMTITERQAAIRNLYLKIILRQELAFDPKYNTBEVVGMSGDVLIQDAAGE 190  
 QY 119 KIALLFQNMSTFSGIAGVLGKWKLTLYLTSTPLMASAACSRLVLSLTKELSAVS 178  
 DB 191 KVGKFIQLVTFLLGFIYAFQGNHLLTYMMAITPLVLAAGVSNVYAKNASLQQAAYA 250  
 QY 179 KAGAAVEVLSIRTVIAFRAOEKELQSRFLNITRYAMFYEPQWLLSCVLVFRYQNL 238  
 DB 251 ESSVVEGTGISIRVASFTEGKAVE-----KYNKSL 283  
 QY 229 KDADFGIKRITIASKVSLSGAVYFPMNGTGLAFWYGTSLINCEGTYITGLVAVFESVI 298  
 DB 284 KSAVSGVREGIAGLGTWVLLFCGYSGLIWGAKLL--LKGYGAVVMVYIFAVL 341  
 QY 299 HSSYICGAAPHEPFIARGAFFHFOVDKPKSIDNFSTAGYKPESEGEFNKNSF 358  
 DB 342 TGSIALGQASPSMAFAAGQAAATKMEFTINKPEIDAVSTTGAKPDIDRGDIEFRDYF 401  
 QY 359 NPSRPSIKIILKGLNLRIKSGETVALVGLSGKSTVYVOLLQRLVDPDDGFIMVENDIR 418  
 DB 402 SYTPRPDEQIRGFSLSIPSGTTVALVQSGSKSTVSLIERFYDPQLGVLIDGVNLK 461  
 QY 419 ALNVRHRYDHIGVSOEPLVFGTTISNNIKYGRDDVTDDEMERARREANAYDIFMEFNK 478  
 DB 462 EFOLRWIRSKIGLVSOEPLVFAASIKENIAGKDNATOEIRAAELMNAKRFIDKMPQG 521  
 QY 479 FNTLVGEKGAOMSGOKORAIARALVRNPKLILIDEATSALDSKSAVOALEK 534  
 DB 522 LDTSVGEHGTOLSGGOKORAIARAIKDPKRLILDEATSALDAESEKIVQEAALDR 577  
 RESULT 15  
 Q9V6I6 PRELIMINARY; PRT; 1279 AA.  
 AC Q9V6I6;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE MDR49 PROTEIN.  
 GN MDR49 OR CG3879.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OK NCBI\_TaxID=7227;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abri J.F., Abpayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borckova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burfils K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellan A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., Moleod M.P., McPherson D.,  
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "the genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 CC EMBL: AE003820; AAF58437.1; -;  
 DR FLYBase: FBgn0004512; Mdr49.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW SEQUENCE 1279 AA; 140075 MW; FDDAD8DC0CE3B812 CRC64;

Query Match 36.68; Score 1002; DB 5; Length 1279;  
 Best Local Similarity 39.28; Pred. No. 1.8e-59;  
 Matches 222; Conservative 95; Mismatches 189; Indels 60; Gaps 6;

QY 1 MIIGLASLVGACLPMPVLYGEMSDNLISGLVOTNTYSEFRLTLTYVG----- 51  
 DB 46 LVVSLVATASAFIFPMIITGFTSLVDRYVGVTSSPAFLPMFGGQQLTNASK 105  
 QY 52 -----IGVAALIFG---YIOISLWITTAAR---QTKRIKQFPHSVLAODIG 92  
 DB 106 ENNQAIIDDATAFGISLVGSVAMEFLITLAIIDLANRIALNQIDRIKRLFEAMLRDIA 165  
 QY 93 WFDSCDIDGELTNRMT-DIDKTSIDGIGKIALLFQNMSTFSIGLAVGLYKKGKLLIYLTST 151  
 DB 166 WYDTRSSGSNFASKMTEDLDLKEIGEKIVIVELIMTFVIGIVSAFVYGMKLTLLVLS 225

QY 152 SPLIMASAAACSRMWTSLTSELSANSKAGAAVEVLISRTIYIAFRAQKELOBSFLN 211  
 DB 226 VPELIATSVARLQGLSLAKELSYSDAANVVEEVSFSGIRTYAFASGQKEKE----- 279  
 QY 212 ITTYAMFYFQWLLSCLVAVRYTONLCKDAKDFGIRKTIASKVSLGAVYFEMNGTYGLAF 271  
 DB 280 -----RQKLLIPENYGRKKGLISGMGNLSWLIITLCMALAI 318  
 QY 272 WYGTSLILN---GEPYTTIGYLAFFSVYHSSYICGAAPPHFETPAIARGAAPHIFOV 327  
 DB 319 WYGTTLILDBRDLDPDRYVTPAVLYVLFVAYMGAQNQNGFASPHVEALAVATAAGQLFNI 378  
 QY 328 IDKPSIDNFSTAGYKRESIEGYEFKNVSPNPSRSITILGMLNRTISGETVALYGL 387  
 DB 379 IDRPSQVDPMDKGNRDEMTAGHTRFEGIRFRYPARPDAVEILGLTVDLVPGQVAFVGA 438  
 QY 388 NGSGKSTVYOLLRLYDPDDGFIVDENDJRALNRHRYRPHIGIVSGEPVLFETTSNNT 447  
 DB 439 SGCGKSTLIQLMQRFYDPPEAGSVKLDGRDRLTNVGLRSQIGVYGEPVLFATTIGENI 498  
 QY 448 KYGRDVTDEMERARAREANAYDIMEFPNKFNTLVGEKGAQMSGGOKRIATARALVRN 507  
 DB 499 RYGRPSATQADIEKARAANCHOEITRLPKGYDTQVEKGAGQISGGOKRIATARALVRQ 558  
 QY 508 PKIILDEATSAIDSESKSAVQAALE 533  
 DB 559 PQVLLDEATSAIDPTSEKRVQSALE 584

Search completed: April 22, 2002, 14:37:10  
 Job time: 262 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:34:22 ; Search time 61.37 Seconds

(without alignments)  
620.395 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592  
Sequence: 1 MIIGLIASLVNAGACLPMLPL.....SEKSAVQAALKEKTPRYSF 514

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

A\_Geneseq\_1101.\*  
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1580	61.3	1276	14	AA1935199
2	1590	61.3	1276	21	AA158188
3	1585	61.1	1281	22	AAE00310
4	1581	61.0	1281	22	AAE00308
5	1581	61.0	1281	22	AAE00309
6	1580.5	61.0	1283	22	AA181065
7	1579	60.9	1280	18	AA144073
8	1579	60.9	1280	21	AA158186
9	1579	60.9	1280	22	AA181959
10	1576	60.8	1275	21	AA178879
11	1576	60.8	1275	22	AA160409

12	1576	60.8	1280	19	AA148997
13	1576	60.8	1280	19	AA148999
14	1576	60.8	1280	22	AA181068
15	1576	60.8	1280	22	AAE00304
16	1576	60.8	1281	22	AAE00303
17	1574	60.7	1280	22	AA104347
18	1572	60.6	1280	22	AA181064
19	1570	60.6	1280	15	AA163624
20	1570	60.6	1280	15	AA148998
21	1570	60.6	1280	19	AA149000
22	1570	60.6	1280	21	AA158187
23	1570	60.6	1280	22	AA181066
24	1570	60.6	1280	22	AAE00306
25	1568	60.5	1280	8	AA170452
26	1567	60.5	1280	11	AA104868
27	1566.5	60.4	1272	21	AA170597
28	1564.5	60.4	1272	21	AA170596
29	1564	60.3	1280	14	AA144297
30	1562.5	60.3	1279	22	AA181067
31	1562.5	60.3	1279	22	AAE00307
32	1559.5	60.2	1276	21	AA158189
33	1553.5	59.9	1279	19	AA180294
34	1050.5	40.5	1275	20	AA182594
35	975.5	37.6	1408	19	AA162871
36	965.5	37.2	1243	21	AA139102
37	965.5	37.2	1286	21	AA139101
38	959	37.0	1349	20	AA116434
39	947	36.5	1205	21	AA139103
40	937	36.1	1334	20	AA102630
41	927.5	35.8	1254	21	AA192173
42	920	35.5	1307	17	AA199255
43	889	34.3	1333	22	AA162495
44	858	33.1	1261	22	AA162493
45	811	31.3	570	21	AA11406

#### ALIGNMENTS

RESULT 1	AA1935199	standard; Protein; 1276 AA.
ID	AA1935199	
XX	AA1935199	
AC	28-JUL-1993	(first entry)
DT	Mouse multidrug resistance protein.	
XX		
DE	mdr gene; Lambda DR11 clone.	
XX		
KW	Mus musculus.	
XX		
OS	US198344-A.	
PN	30-MAR-1993.	
XX		
PD	15-JUL-1986;	86US-0885951.
XX		
PF	15-JUL-1986;	86US-0885951.
XX		
PR	15-JUL-1986;	86US-0885951.
PR	06-FEB-1991;	91US-0652311.
XX		
PA	(MAST) MASSACHUSETTS INST TECHNOLOGY.	
XX		
PI	Croop JM, Gros P, Housman DE;	
XX		
XX	WPI: 1993-126077/15.	
DR	N-PSDB; AAQ38950.	
XX		
XX	DNA sequence which confers multi-drug resistance on sensitive	
PT	mammalian cells - used to preserve bone marrow cells during	
PT	chemotherapy to prevent infection	

Wild-type human P  
-C-terminal single  
Dog P-glycoprotein  
Dog P-glycoprotein  
Dog (PGP) P-glycoprotein  
Human multidrug re  
Cynomolgous monke  
Human P-glycoprote  
N-terminal single  
Human P-glycoprote  
Human G185V mutant  
Human P-glycoprote  
Human P-glycoprote  
Sequence encoded b  
Protein encoded b  
Rat multidrug resi  
Rat multidrug resi  
Sequence encoded b  
Human P-glycoprote  
Human P-glycoprote  
Human P-glycoprote  
Mutated human P-g  
H. contortus PGP-A  
Multiple drug resi  
Arabidopsis thalia  
Arabidopsis thalia  
Multiple drug resi  
Arabidopsis thalia  
Arabidopsis thalia  
Protein encoded by  
A. thaliana ATPAC,  
Aspergillus flavus  
P. chrysogenum ABC  
P. chrysogenum ABC  
Arabidopsis thalia



PS Disclosure; Fig 8; 22pp; English.

XX The lambda Dr11 mdr gene was isolated as described in AA038950.  
 CC The amino acid sequence deduced from its open reading frame has  
 CC potential glycosylation sites and is likely to be a membrane  
 CC protein. The encoded protein appears to contain (starting at the  
 CC 5'-end) three hydrophobic domains, a neutral domain, three  
 CC additional hydrophobic domains and a second neutral domain. There  
 CC are ATP-binding sites within the neutral domains. It is  
 CC hypothesised that increased synthesis of the membrane protein may  
 CC alter membrane physiology to cause decreased intracellular drug  
 CC accumulation.

XX Sequence 1276 AA;

Query Match 61.3%; Score 1590; DB 14; Length 1276;  
 Best Local Similarity 57.7%; Pred. No. 8.7e-139;  
 Matches 307; Conservative 99; Mismatches 94; Indels 32; Gaps 5;

QY 1 MILGILASLVNGACLPMPVLYGEMSDNLI-----SGCLVQNTN----- 39  
 DB 50 mlglglaaingtlplmllyfgmtddsfkkaasilpsitngsgpnstllissnslee 109  
 QY 40 ---YSEFRLTYVYGIVAAALFGYIOISLWITTAARQTKRIKQFPHSVLAQDIGWDS 96  
 DB 110 malayay-----yvtgigagvavilvaylvsfwclaaagrqhkrqgfahaimqelgwtdv 164  
 QY 97 CDIGELNTRMT-DIDKISDGIQDKTALLFQNMSTFSGIAGLVGKWLTVLTSTSPLI 155  
 DB 165 hdvgeintnrltdvskindgiskgmftgslttlagflfigisgkwltvlavsp11 224  
 QY 156 MASAACSRMWISLTSKLSAVSKAGAAVEYLSIRTVIAFRAOEKLORYTONLKAOK 215  
 DB 225 glssalwakvltstfnkelgagayakagaveevlaaitrviaifgsgqkelerynkleeak 284  
 QY 216 DPGIRTRTASKVSICAVVFPNMGTYGLAFWGTSLILNGEGYTGTVLAVFVSYHSSY 275  
 DB 285 nvglkkaitsisglayllvyasayalafwytslvlsne--ysigevltvffsllgtf 342  
 QY 276 CIGAAVPHFETFAIRGAAPHIFOVYIDKKPSIDNSTAGYKPESTIEGTVEKKNVSFNPS 335  
 DB 343 signlaphiaaefanergaatefklidnepsidstfkyrpsdimgnleknvhtnyps 402  
 QY 336 RPSIKILKGLMLRIKSGEVALVGLNGSGKSTVYOLQRLYDPPDGFIMVENDIRALN 395  
 DB 403 rsevgilkglnlkvksqgtvalvngsgckstvtgltmqrllydpiegavvsidgqdlrtlnv 462  
 QY 396 RHYRDHIGVSOEPLYLEFITTSSNNIKYGRDVTDEEMERAAEANAADYFIMEFPKENTL 455  
 DB 463 tylteligvvsqepvlfatlaenirgyredvtmdelakavkeanaaydfimklphqfdtl 522  
 QY 456 VGEKGAGWGGOKORIAFARALVRNPKILLIDEATSALDSEKSAVQAALXK 507  
 DB 523 vgergaqisggqkqrialaralvrnpxkllilideatsaldeeaavgaaldk 574

RESULT 2

AA58188 ID AA58188 standard; Protein: 1276 AA.

XX AA58188;

XX 14-MAR-2000 (first entry)

XX Murine multidrug resistance-1 (MDR-1).

XX Multidrug resistance; MDR-1; P-glycoprotein;

XX transmembrane efflux pump; haematopoietic stem cell; transduction;

KM bone marrow transplantation; chemotherapy; radiation therapy; cancer;

KM gene therapy; gene replacement; genetic defect; thalassemia;

KM Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;

KM cytokine.

XX Mus sp.

OS W09961589-A2.

XX 02-DEC-1999.

XX 27-MAY-1999; 99WO-US11825.

XX 28-MAY-1998; 98US-0066988.

XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX Sorrentino B, Bunting K;

XX WPI: 2000-072615/06.

XX N-PSDB; AA249334.

PT Ex vivo expansion of hematopoietic stem cells transduced with a  
 sequence encoding human multidrug resistance-1, used for bone marrow  
 transplantation -

PS Disclosure; Page 93-101; 113pp; English.

XX This sequence represents murine multidrug resistance  
 CC protein MDR-1. MDR-1 is a transmembrane efflux pump,  
 CC responsible for the export of drugs from cells, particularly  
 CC cancer cells. The invention relates to transducing hematopoietic  
 CC stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified hematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC hematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from hematopoietic stem cells, e.g. thalassemia,  
 CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in hematopoietic stem cells.  
 CC Haematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC hematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.

XX Sequence 1276 AA;

Query Match 61.3%; Score 1590; DB 21; Length 1276;  
 Best Local Similarity 57.7%; Pred. No. 8.7e-139;  
 Matches 307; Conservative 99; Mismatches 94; Indels 32; Gaps 5;

QY 1 MILGILASLVNGACLPMPVLYGEMSDNLI-----SGCLVQNTN----- 39  
 DB 50 mlglglaaingtlplmllyfgmtddsfkkaasilpsitngsgpnstllissnslee 109  
 QY 40 ---YSEFRLTYVYGIVAAALFGYIOISLWITTAARQTKRIKQFPHSVLAQDIGWDS 96  
 DB 110 malayay-----yvtgigagvavilvaylvsfwclaaagrqhkrqgfahaimqelgwtdv 164  
 QY 97 CDIGELNTRMT-DIDKISDGIQDKTALLFQNMSTFSGIAGLVGKWLTVLTSTSPLI 155  
 DB 165 hdvgeintnrltdvskindgiskgmftgslttlagflfigisgkwltvlavsp11 224  
 QY 156 MASAACSRMWISLTSKLSAVSKAGAAVEYLSIRTVIAFRAOEKLORYTONLKAOK 215  
 DB 225 glssalwakvltstfnkelgagayakagaveevlaaitrviaifgsgqkelerynkleeak 284  
 QY 216 DPGIRTRTASKVSICAVVFPNMGTYGLAFWGTSLILNGEGYTGTVLAVFVSYHSSY 275  
 DB 285 nvglkkaitsisglayllvyasayalafwytslvlsne--ysigevltvffsllgtf 342  
 QY 276 CIGAAVPHFETFAIRGAAPHIFOVYIDKKPSIDNSTAGYKPESTIEGTVEKKNVSFNPS 335

```

Db 343 siglhapnieafanargaefefkldnepsidsfstckgyypsimgnlefknhfnyps 402
Oy 336 RPSKILKGLNLRKSGETVALVGLNSGKSTVYQLRLYDPDGFIMVENDIRALNV 395
Db 403 rsevgllkglnlkvsqglvalvgnsgcgkstlvqmqrllydpliegvsldgqdlrtlnv 462
Oy 396 RHYRDHGVVSOEPVLFGTTSSNNIKYGRDQVTDDEMERARAEANAYDFIMEFPKFTL 455
Db 463 ryltelitgvsqepvlfatcltaenlrygreuvmdelekaveanaydfimklnpqtcll 522
Oy 456 VGEKAGMSGGOKORAIARALVBNPKILIDEATSALDSEKSAVQAALRK 507
Db 523 vgergaqslsgqkqrlatlaralvrnpklllldeatsaldteseeavqaalrk 574

RESULT 3
AAE00310
ID AAE00310 standard; Protein: 1281 AA.
XX
AC AAE00310;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype D).
XX
KM Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KM MDR1; drug bioavailability; transgenic animal; genetic model.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT MISC-difference 25 /note= "asn of Genotypec substituted by Lys"
FT MISC-difference 197 /note= "His of Genotypec substituted by Gln"
FT MISC-difference 329 /note= "Ser of Genotypec substituted by Thr"
FT MISC-difference 1148 /note= "Met of Genotypec substituted by Val"
FT
XX
PN W0200123540-A2.
XX
PD 05-APR-2001.
XX
PE 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI: 2001-235373/24.
DR N-PSDB; AAD03506.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors.
XX
XX Claim 17; Page 108-110; 111pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense-oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
```

```

CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant.
CC This sequence is also referred as Genotype D protein. The PGP
CC enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 1281 AA;
XX
Query Match 61.1%; Score 1585; DB 22; Length 1281;
Best Local Similarity 57.5%; Pred. No. 2.5e-138;
Matches 304; Conservative 97; Mismatches 104; Indels 24; Gaps 4;
Oy 1 MILGLASLVNGACPLPLVLGEMSDNLISGCL-----VQNTYSFR--- 44
Db 51 mlygtmaailhgaalpjmmllvfgmtdsfanaglsrntkfpvlneslntnqfhlhle 110
Oy 45 -----LFLYVGIGVAAALIFGYIOISLWITTAAROTKRIRKOFHSLAODIGMFDSCI 99
Db 111 eemltayayysgigavlvaaayiqvswcilaagqllkirkqfhaimgelgvfdvndv 170
Oy 100 GELNTRMT-DIDKISDSDIGDKIALLFQNMSPFSIGLAVGLVKGWKLVTLSPLIMAS 158
Db 171 gelntrltdvsklnegldgkigmffgsiatffgfvlgtrgwkltlvllalspvlgls 230
Oy 159 AAACSRWVLSLTSKELSAVSKAGAAVEVLSIRVINFRAQEKLOXTQNLDAKQFG 218
Db 231 aaiwakllsftfdeklldayakagaaevlaaitrlvalafigqkkeleryknleakgig 290
Oy 219 IKRTASKVSIGAVYFFNGNYGGLAFWYGTSLINCEPGYITGFVLAVFVSIHSSYICG 278
Db 291 ikkatamtsigaeflllyasayalafwygtalvsse--ytlgqyltffvsligafsig 348
Oy 279 AAVPHETFAIRGAAPHIFQVIDKRPISIDNESTAGYKRESIEGTEKKNVSFNYSRPS 338
Db 349 gaspsieafanargaeyefkldnkpidsyskghpdklgnlefknhfysyrke 408
Oy 339 IKILKGLNLRKSGETVALVGLNSGKSTVYQLRLYDPDGFIMVENDIRALNVHY 398
Db 409 vkllkglnlkvsqglvalvgnsgcgkstlvqmqrllydpliegvsldgqdlrtlnvzl 468
Oy 399 RDHIGVSOEPVLFGTTSSNNIKYGRDQVTDDEMERARAEANAYDFIMEFPKFTLGE 458
Db 469 reitgvsqepvlfatcltaenlrygreuvmdelekaveanaydfimklnpqtcllvg 528
Oy 459 KGAOMSGGOKORAIARALVBNPKILIDEATSALDSEKSAVQAALRK 507
Db 529 rgaqslsgqkqrlatlaralvrnpklllldeatsaldteseeavqaalrk 577

RESULT 4
AAE00308
ID AAE00308 standard; Protein: 1281 AA.
XX
AC AAE00308;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype A).
XX
KM Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KM MDR1; drug bioavailability; transgenic animal; genetic model.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT MISC-difference 197 /note= "His of Genotypec substituted by Gln"
FT
XX
PN W0200123540-A2.
```

XX	05-APR-2001.
XX	
PF	28-SEP-2000; 2000OWO-US26767.
XX	
PR	28-SEP-1999; 99US-0156510.
XX	
PA	(GENT-) GENEST CORP.
XX	
PI	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX	
DR	WPI: 2001.235373/24.
DR	N-PSDB: AAD03504.
XX	
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful for determining the bioavailability of drugs and for screening for dog PGP inhibitors -
XX	
PS	Claim 17; Page 91-93; 11pp; English.
CC	
CC	The invention relates to dog P-glycoprotein (PGP) also referred
CC	as multidrug transporter (MDR1) and nucleic acids encoding them.
CC	The invention also includes fragments and biologically functional
CC	variants of dog P-glycoprotein. PGP and their nucleic acids are
CC	useful for determining the bioavailability of drugs and for
CC	screening PGP inhibitors. They are useful for the diagnosis and
CC	treatment of conditions characterised by PGP activity, by
CC	reducing or increasing PGP activity in a cell. PGP nucleic acids
CC	are used as oligonucleotide probes. Complements of PGP nucleic
CC	acids are useful as antisense oligonucleotides, to induce a PGP
CC	'knockout' phenotype. They are used to prepare a non-human
CC	transgenic animal, which are valuable as genetic models for
CC	human diseases.
CC	The present sequence is dog P-glycoprotein (PGP) allelic variant.
CC	This sequence is also referred as Genotype A protein. The PGP
CC	enzyme functions as an efflux pump exporting small molecules
CC	across the cell membrane. This enzyme is a member of the ABC
CC	transporter family.
XX	
SO	Sequence 1281 AA;
	Query Match 61.0%; Score 1581; DB 22; Length 1281;
	Best Local Similarity 57.3%; Pred. No. 6e-138;
	Matches 303; Conservative 98; Mismatches 104; Indels 24; Gaps
OY	1 MILGLAIVNGACPLMPVLVGENSDNLISCL-----VQTNYSFR--- 44
Db	51 mlvgtmaailngaaipflmilyfgmtltdsfanagisnrnkftfyllnesitnltqhinhle 110
OY	45 -----LTLTYVGIGAGVALIFGYIOISLWITRAOTKRIRKOFNFSSVAODIGWFDSCDI 99
Db	111 eemtlayayygsigsgqvvaayiygvsfciaagrilkirkffinaimqegwfdvndv 170
OY	100 GELNTRMT-DIDIKTSDGIGDKITALLFQNMFSTFSGILAVGLVKWKRLTVTTISTSPILMAS 158
Db	171 gelntrlcdsvskinegjdkgifmgfsiatftfigvfgrwxklltvlaispvlgjs 230
OY	159 AACGRMVIISLTSKRLSAYSKAGANAEEVLSIRIVIAFRAOEKELORTOLMKAKDPG 218
Db	231 aaawklisfstfdkellayakagaavaeevlaalrtviafgygqkelerynkhleakgig 290
OY	219 IKRTASVSVSGAAVFPMFGVTGLAFWVGTSIIUNGEQYTGTGVLAVFVSYTHSYCYG 278
Db	291 ikkaiatnisigaafilliyaasalafwygstslisse--ysigqyltvtfvisligafsigs 348
OY	279 AAVPHETFAIRGAAPHIFOVIDRKPSINDNSTAGYRPESIEGTVEFKNSFNPNPSRPS 338
Db	349 gaspieaeafanaargaayelfikidhkpsidsyksghpnmdlnkgnlefnnvfysprke 408
OY	339 IKILKGLWRISKGETVALVGLNGSGKSTVOOLLRLDPPDDGFLMVDENDIRALNVRYH 398
Db	409 vkilkgknlhkyagsgtvalavngmscgksktlvjlmrplydptlgnawciogdqdlrtlnvrhl 468

Oy		399	RDHIGVSOAEVLFGCTISNNIKVGDDVDYDEMEFAEAENANNOYFIMEFPKFTLIGE	458
Oy			:                       :   :   :   :   :   :   :   :   :	
Dd		469	reitgvsvgepvitatiaenitrygrenvmdolekaveanaydfimklipkfticlviqe	528
Oy		459	KGAOMSGGOKORIRAIARALVRNPKNRIILDEATSAIDSESKSAVOALEK	507
Dd		529	rqaqslsggqqriralatarylrrnpkhilldeatsalidseavvyvaldk	577
RESULT	5			
ID	AAE00309			
XX	AAE00309 standard; Protein; 1281 AA.			
XX	AAE00309;			
DT				
XX	13-JUN-2001 (first entry)			
DE				
XX	Dog P-glycoprotein (PGP) allelic variant (Genotype B).			
KW				
XX	Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;			
XX	MDR1; drug bioavailability; transgenic animal; genetic model.			
OS				
XX	Canis familiaris.			
FH				
FT	Key	Location/Qualifiers		
FT	Misc-difference 25	/note= "asn of Genotynec substituted by lys"		
FT	Misc-difference 197	/note= "His of Genotynec substituted by Gln"		
XX				
PX	WO200123540-A2.			
PD				
XX	05-APR-2001.			
PF				
XX	28-SEP-2000; 2000WO-US26767.			
PR				
XX	28-SEP-1999; 99US-0156510.			
XX				
PA	(GENT-) GENTEST CORP.			
PI				
XX	Stocker PJ, Steimel-Crespi DT, Crespi CL, Relif TC, Patten CJ;			
XX				
DR	WPI; 2001-235373/24.			
N-PSDB; AAD03505.				
PT				
XX	New dog P-glycoproteins (PGP) and their encoding nucleic acids; useful			
XX	for determining the bioavailability of drugs and for screening for dog			
XX	PGP inhibitors -			
Claim 17; Page 99-102; 11pp; English.				
The invention relates to dog P-glycoprotein (PGP) also referred				
as multiliding transporter (MDRL) and nucleic acids encoding them.				
The invention also includes fragments and biologically functional				
variants of dog P-glycoprotein. PGP and their nucleic acids are				
useful for determining the bioavailability of drugs and for				
screening PGP inhibitors. They are useful for the diagnosis and				
treatment of conditions characterised by PGP activity; by				
reducing or increasing PGP activity in a cell. PGP nucleic acids				
are used as oligonucleotide probes. Complements of PGP nucleic				
'acids are useful as antisense oligonucleotides, to induce a PGP				
'knockout' phenotype. They are used to prepare a non-human				
transgenic animal, which are valuable as genetic models for				
human diseases.				
The present sequence is dog P-glycoprotein (PGP) allelic variant.				
This sequence is also referred as genotype B protein. The PGP				
enzyme functions as an efflux pump exporting small molecules				
across the cell membrane. This enzyme is a member of the ABC				
transporter family.				
Sequence 1281 AA;				

Query Match 61.0%; Score 1581; DB 22; Length 1281;  
 Best Local Similarity 57.3%; Pred. No. 6e-138;  
 Matches 303; Conservative 98; Mismatches 104; Indels 24; Gaps 4;

QY 1 MILGLASLVNACPLPLVLVGENSDNLISGCL-----VQNTYSEFR--- 44  
 DB 51 mlvgmaailhgaalplmvlvfgmntdsfanagisrntkfpvlnesltnqhtfinhle 110  
 QY 45 -----LTIYVYGIVAAALIFGCIQISLMIITRAOTKRIKROFHSVLAODIGMFDSCDI 99  
 DB 111 eemttayayysqigagvlvaaylqysfwcslaagrqllkrlqfifhalmrgeqywdvndv 170  
 QY 100 GELNTRMT-DIDKISDGIKIDKIALLFQNMSTFSGISGLAVLWGMKLTVTSTSPILMAS 158  
 DB 171 gelntrldtdvskinegysgdkimgfsgiatffgfygftcrgwkltlvllaipvlqls 230  
 QY 159 AAACSRMVISLTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELORYTONLKDAPFG 218  
 DB 231 aalwakiisftckellayakagavaeevlairtviafsgqkkelerynknleakgig 290  
 QY 219 IKRTIASKVSIGAVYFPMNGTYGLAFWYGTSLJLNGEPYITGTVLAVFVSYSYCI 278  
 DB 291 ikkaltanisiqaalliyasayalaftwysltvisse--yslgylytfvffsvllgafsig 348  
 QY 279 AAVPHFEFTALARGAAPHIFQVIDKPPSIDNFTAGYKPESTEGTVEPKNVSFNPSRPS 338  
 DB 349 qaspsieafanargaaayelfkldmkpsidsysgshkpdnklgnlftknvlfyspsrke 408  
 QY 339 IKILGLMLRIKSGEFTVALVGLNSGKSTVVOQLORLYDPDDGFTMVDENDIRALNVRY 398  
 DB 409 vkllkglnlkvsgqglvalvngsgcstvtgqlmgrylvpdpcgmcidgqdlrtlnvrl 468  
 QY 339 RDHIGVYQOEVLFGTTSSNNIKYGRDVTDEEMERARERANAYDFIMEFPNKFTLVGE 458  
 DB 469 reltvgvsgqglvaltlaenlrygrenvmdelkavkeanaaydfimkplpkihtlyge 528  
 QY 459 KCAQMSGGOKORIAFARALVRNPKLILIDEATSAIDSESKSAVOALEK 507  
 DB 529 rgaqsgsqgkqrlaalaralvrnprkllldeatsaideseavvqalidk 577

RESULT 6  
 AAB81065  
 ID AAB81065 standard; Protein: 1283 AA.

AC AAB81065;  
 XX  
 DT 25-JUN-2001 (first entry)  
 XX  
 DE Cynomologous monkey P-glycoprotein variant 2.  
 KW Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDRI;  
 KM efflux pump.  
 XX  
 OS Macaca fascicularis.  
 XX  
 EH Key Location/Qualifiers  
 FT Misc-difference 93..95 "An additional 3 amino acids are present compared  
 FT to PGP variant AAB81064"  
 XX  
 XX MO200123565-A1.  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000OWO-US26592.  
 XX  
 PR 28-SEP-1999; 99US-0156921.  
 XX 12-OCT-1999; 99US-0158818.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 PI Stocker PJ, Stelmel-Crespi DT, Crespi CL;

XX  
 DR WPI: 2001-316136/33.  
 DR N-PSDB: AAF86128.  
 XX  
 PT Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein  
 PT (PGP) and homologous PGP polypeptides are useful for predicting  
 PT bioavailability of compound and increasing PGP transporter activity in  
 PT cell  
 XX  
 PS Claim 9; Page 65-68; 8app; English.  
 XX  
 CC This invention relates to a polynucleotide sequence encoding a  
 CC cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the  
 CC PGP protein. PGP, also known as multidrug transporter, MDRI is a member  
 CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
 CC exporting small molecules across the cell membrane. The invention  
 CC includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence  
 CC and protein, and also that of an allelic variant. The PGP polynucleotide  
 CC sequence is useful for increasing PGP transporter activity in a cell.  
 CC Antisense sequences of the cDNA are useful for inhibiting PGP transport  
 CC activity in a mammalian cell. They may also be used for increasing the  
 CC bioavailability of a drug. The present sequence represents the  
 CC cynomologous monkey P-glycoprotein variant 2. The protein has an  
 CC additional 3 amino acids when compared to PGP variant 1 (AAB81065).  
 CC  
 SQ Sequence 1283 AA;

Query Match 61.0%; Score 1580.5; DB 22; Length 1283;  
 Best Local Similarity 57.7%; Pred. No. 6.7e-138;  
 Matches 306; Conservative 96; Mismatches 103; Indels 25; Gaps 4;

QY 1 MILGLASLVNACPLPLVLVGENSDNLIS-----GCLVQNTYS----- 41  
 DB 51 mvvgfialahhgaqllplmvlvfgmdtclftanagnlqdlgallfntnssnlctvypmnl 110  
 QY 42 ---FRLTLTYVYGIVAAALIFGCIQISLMIITRAOTKRIKROFHSVLAODIGMFDSCD 98  
 DB 111 eemttayayysqigagvlvaaylqysfwcslaagrqllkrlqfifhalmrgeqywdvnd 170  
 QY 99 IGEINTRMT-DIDKISDGIKIDKIALLFQNMSTFSGISGLAVLWGMKLTVTSTSPILMA 157  
 DB 171 vgelntrltdvskinegysgdkimgfsgiatffgfygftcrgwkltlvllaipvlql 230  
 QY 158 SAAACSRMVISLTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELORYTONLKDAPF 217  
 DB 231 saavwakliisftckellayakagavaeevlairtviafsgqkkelerynknleakxrl 290  
 QY 218 GIKRTIASKVSIGAVYFPMNGTYGLAFWYGTSLJLNGEPYITGTVLAVFVSYSYCI 277  
 DB 291 gikaltanisiqaalliyasayalaftwysltviske--yslgylytfvffsvllgafsv 348  
 QY 278 GAAPHFEFTALARGAAPHIFQVIDKPPSIDNFTAGYKPESTEGTVEPKNVSFNPSRPS 337  
 DB 349 qaspsieafanargaaayelfkldmkpsidsysgshkpdnklgnlftknvlfyspsrk 408  
 QY 338 SIKILGLMLRIKSGEFTVALVGLNSGKSTVVOQLORLYDPDDGFTMVDENDIRALNVRY 397  
 DB 409 evkllkglnlkvsgqglvalvngsgcstvtgqlmgrylvpdpcgmcidgqdlrtlnvrl 468  
 QY 398 VRDHTGVYQOEVLFGTTSSNNIKYGRDVTDEEMERARERANAYDFIMEFPNKFTLVG 457  
 DB 469 lreilvgvsgqglvaltlaenlrygrenvmdelkavkeanaaydfimkplpkihtlyge 528  
 QY 458 EKAQMSGGOKORIAFARALVRNPKLILIDEATSAIDSESKSAVOALEK 507  
 DB 529 ergaqsgsqgkqrlaalaralvrnprkllldeatsaideseavvqalidk 578

RESULT 7  
 AAM44073  
 ID AAM44073 standard; protein: 1280 AA.

AC	AAW44073:	
AD		
AE		
AF	26-JUN-1998	(first entry)
AG		
AH	Human multidrug resistance P-glycoprotein MDRL.	
AI		
AJ	Human; multidrug resistance P-glycoprotein; MDRL; prokaryotic homologue	
AK		
AL		
AM	Homo sapiens.	
AN		
AO		
AP	Key	Location/Qualifiers
AQ	Region	1..640
AR	Region	/note="MDRL-N from Fig 1"
AS	Region	641..1280
AT		/note="MDRL-C from Fig 1"
AX		
AY	W09740160-A1.	
AZ		
BA	30-OCT-1997.	
BB		
BC	24-APR-1997;	97MO-N000216.
BD		
BE	24-APR-1996;	96EP-0201094.
BF		
BG	(UYGR-) RIJKSUNIV GRONINGEN.	
BH		
BI	Bohluus H, Konings WN, Van Veen HW, Venema K;	
BJ	WPI: 1997-535844/49.	
BK		
BL	Prokaryotic homologue of human multiple drug resistance protein -	
BM	used to screen for compounds that inhibit, or avoid, drug resistance	
BN		
BO	Claim 10; Fig 1; 35pp; English.	
BP		
BQ		
BR		
BS		
BT		
BU		
BV		
BW		
BX		
BY		
BZ		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		

[illegible]

```

Db      231  vwa klss ftdkellayagagavaeevllaatrtvtafggqkkelerynkleekrtigk 290
Qy      221  RTIASVSLGAVYFFPMNTGYDLAAWYGSLLNPEPGYTGTVLAVFFSYTHSSYCGAA 280
Db      291  ka ltn slsiga fl llyasayalaktwylctlylsg -ys lq qv ltv ltfv lly lga fsgqa 348
Qy      281  VPHEFTFAIAGAAFHFLFQVLDKKRPSIDNFESTAQYKPSIEGTVEFNFNFNFPNPSRSIK 340
Db      349  spslaefanargaaeyefkrlidnkp ridsy skgnhkdnlgnlefnvfnfsyrskv 408
Qy      341  ILKGLNLTKSGERYVALYGLGSGSKSTYVOLLQYLVPPDGFIMVDENDIRALNRYIRD 400
Db      409  llkglnlkvsgsqvta lvgnsgskstcvlqmq llylpcteamsvdgdrltlvrlf lre 468
Qy      401  HIGVVSQEPVLEFGTTISNNIKRYGDDVDYDEEMERAAAEANAYDFIMFPFKRNTLYGK 460
Db      469  llgvvsqevp lfa tlaen lryfenvymde lka veeanaydf lmk lpk fcd lvgery 528
Qy      461  AQMSGGCKORAIARALVARNPKTLLDEATGALDSSEKSNVAQALEK 507
Db      559  aqlsgqgqqr lalalaa lvm pkl lllldeatsaldeeavvayqald 575

```

	RESULT	8
ID	AAV58186	
XX	AAV58186 standard; Protein: 1280 AA.	
AC	AAV58186;	
XX		
DT	14-MAR-2000 (first entry)	
DE		
XX	Human wild-type multidrug resistance-1 (MDR-1) protein.	
KW	Multidrug resistance; MDR-1; P-glycoprotein;	
KW	transmembrane efflux pump; haematopoietic stem cell; transduction;	
KW	bone marrow transplantation; chemotherapy; radiation therapy; cancer	
KW	gene therapy; gene replacement; genetic defect; thalassemia;	
KW	Gancher's disease; sickle cell anaemia; leukaemia; ex vivo expansio	
KX	cyclokin; wild-type.	
OS	Homo sapiens.	
FH	Key Location/Qualifiers	
FT	Misc-difference 185	
FT	/note= "This residue is Val in a mutant MDR-1	
FT	(AAV58187)"	
PN	WO9961589-A2.	
PD	02-DEC-1999 .	
PD		
PF	27-MAY-1999; 99WO-US11825.	
PR	28-MAY-1998; 98US-0086988.	
PA	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
PI	Sorrentino B, Bunting K;	
PI		
DR	WPI: 2000-072615/06.	
DR	N-PSDB; AAZ49332.	
PT	Ex vivo expansion of hematopoietic stem cells transduced with a	
PT	sequence encoding human multidrug resistance-1, used for bone marrow	
PT	transplantation -	
PS	Claim 10; Page 71-79; 113pp; English.	
CC	This sequence represents human wild-type multidrug	
CC	resistance protein MDR-1. MDR-1 is a transmembrane efflux	
CC	pump, responsible for the export of drugs from certain	
CC	cells, particularly cancer cells. Wild-type MDR-1 shows	
CC	increased resistance to etoposide and decreased resistance to vinca	

CC alkalooids compared with a mutant form (AAV58187) where the Gly at  
CC position 185 is replaced by Val. The invention relates to transducing  
CC hematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified hematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC hematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from hematopoietic stem cells, e.g., thalassemia,  
CC Gaucher's disease, sickle cell anemia or leukemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in hematopoietic stem cells.  
CC Hematopoietic stem cells that express MDR-1 will be protected against  
CC chemotherapeutic agents, so can be engrafted while the patient is  
CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells  
CC provides sufficient cells to permit standard biochemical analysis.  
CC Overexpression of MDR-1 allows cytokine-driven expansion of  
CC hematopoietic stem cells by at least 10-fold compared with a maximum  
CC of 4-fold in known procedures.

XX Sequence 1280 AA:

Query Match 60.9%; Score 1579; DB 21; Length 1280;  
Best Local Similarity 58.1%; Pred. No. 9.2e-138;  
Matches 306; Conservative 100; Mismatches 99; Indels 22; Gaps 5;

OY 1 MILGILSLVNGACLPMLPLVLEMSD-----NISGCLVQT--NTVSF----- 43  
DB 51 mvvglaaillngaglpmlmvlfgemtdlfanagnledlmsntlrnsdindgfmmleed 110  
OY 44 --RLTLVVGIGVALIFGYIQISLWITTAARQTRIKRQFPHSVLAODIGFDSCDIGE 101  
DB 111 mtryayysgagvlyvaayiqvswcfaagrqjhkirkqffhaimrgeigvfdvhdvge 170  
OY 102 LNTWMT-DIDKISGIGDKIALPQNMSTFSGIAGLVKWKLTVLTSPLIMASAA 160  
DB 171 Intrldvskinegigdklqgmftgsmatffgtgfvgrgwklvlvllaalspvglsaa 230  
OY 161 ACSRMVSLTSKELSAVSKAGAAVEVLSIRTVAFPAQEKELORYQNOLKDAKDFIRK 220  
DB 231 wvakllsftckellayakagavaeevllaalrtvialatgqgkkelerynklnleearrrigk 290  
OY 221 RTIASKSLAGVAFPMNGTYGLAFWYGTSLINGEGYITGTVLAFPSVSHSVCIGAA 280  
DB 291 kaitanlsigaaflliyasyalaftwygtllvlsge--ysigqvlvtfsvlgaatsvga 348  
OY 281 VPHETFAIARGAAPHIFQVIDKKPSIDNFSTAGYKPSIEGTVFEKKNVSFNPSPRSIK 340  
DB 349 spsleaafanargaayelkikldnkpssidsyskshpdklqgnlefnvhsysprkev 408  
OY 341 ILKGLNLRKSGEYVALVGLNGSGKSTVVOQLQRLYDPDQGFIMVDENDIRALNRYHRD 400  
DB 409 ilkglnlkvsgqetvalvngscgkstvtqlmqrllydptegmvsvdgqdditlrnfrlre 468  
OY 401 HIGVSOEPVLEFQTTISNNIKYGRDQVDEMERARAEANAYDFIMEPPKNFNTLVGEKG 460  
DB 469 ilgvsgqetvalvngscgkstvtqlmqrllydptegmvsvdgqdditlrnfrlre 528  
OY 461 AQMSGGQKRIARALVARNPKIILDEATSALESKSAVOALEK 507  
DB 529 aqlsggqetvalvngscgkstvtqlmqrllydptegmvsvdgqdditlrnfrlre 575

RESULT 9  
AAB81959 ID AAB81959 standard; protein: 1280 AA.

XX AAB81959:

XX 02-JUL-2001 (first entry)

XX Human MDRL.

XX

KW Human; MDRL; multi-drug resistance; cholesterol uptake;  
KW hypercholesterolaemia; hypocholesterolaemia; atherosclerosis;  
KW coronary artery disease; cerebral vascular disease.

XX Homo sapiens.

XX W0200121762-A2.

XX 29-MAR-2001.

XX 22-SEP-2000; 2000WO-US26099.

XX 23-SEP-1999; 99US-0155819.

XX (BARN-) BARNES-JEWISH HOSPITAL.

XX Stenson WF, Tessonner T;

XX WPI; 2001-328100/34.

PT Modulating cellular cholesterol uptake for treating hyper or  
PT hypocholesterolemia. Involves administering an agent that inhibits or  
PT increases the expression of multiple drug resistance-1.  
PS Example 1; Page 43-47; 47pp; English.

CC The present invention describes a method of modulating cholesterol uptake  
CC in cells, involving administering an agent capable of inhibiting the  
CC multi-drug resistance protein MDRL. This is useful in the prevention and  
CC treatment of cholesterol-related diseases, including  
CC hypercholesterolaemia, hypocholesterolaemia, atherosclerosis, coronary  
CC artery disease and cerebral vascular diseases. The present sequence is  
CC the human MDRL protein.

XX Sequence 1280 AA:

Query Match 60.9%; Score 1579; DB 22; Length 1280;  
Best Local Similarity 58.1%; Pred. No. 9.2e-138;  
Matches 306; Conservative 100; Mismatches 99; Indels 22; Gaps 5;

OY 1 MILGILSLVNGACLPMLPLVLEMSD-----NISGCLVQT--NTVSF----- 43  
DB 51 mvvglaaillngaglpmlmvlfgemtdlfanagnledlmsntlrnsdindgfmmleed 110  
OY 44 --RLTLVVGIGVALIFGYIQISLWITTAARQTRIKRQFPHSVLAODIGFDSCDIGE 101  
DB 111 mtryayysgagvlyvaayiqvswcfaagrqjhkirkqffhaimrgeigvfdvhdvge 170  
OY 102 LNTWMT-DIDKISGIGDKIALPQNMSTFSGIAGLVKWKLTVLTSPLIMASAA 160  
DB 171 Intrldvskinegigdklqgmftgsmatffgtgfvgrgwklvlvllaalspvglsaa 230  
OY 161 ACSRMVSLTSKELSAVSKAGAAVEVLSIRTVAFPAQEKELORYQNOLKDAKDFIRK 220  
DB 231 wvakllsftckellayakagavaeevllaalrtvialatgqgkkelerynklnleearrrigk 290  
OY 281 RTIASKSLAGVAFPMNGTYGLAFWYGTSLINGEGYITGTVLAFPSVSHSVCIGAA 280  
DB 291 kaitanlsigaaflliyasyalaftwygtllvlsge--ysigqvlvtfsvlgaatsvga 348  
OY 221 VPHETFAIARGAAPHIFQVIDKKPSIDNFSTAGYKPSIEGTVFEKKNVSFNPSPRSIK 340  
DB 281 VPHETFAIARGAAPHIFQVIDKKPSIDNFSTAGYKPSIEGTVFEKKNVSFNPSPRSIK 340  
DB 349 spsleaafanargaayelkikldnkpssidsyskshpdklqgnlefnvhsysprkev 408  
OY 341 ILKGLNLRKSGEYVALVGLNGSGKSTVVOQLQRLYDPDQGFIMVDENDIRALNRYHRD 400  
DB 409 ilkglnlkvsgqetvalvngscgkstvtqlmqrllydptegmvsvdgqdditlrnfrlre 468  
OY 401 HIGVSOEPVLEFQTTISNNIKYGRDQVDEMERARAEANAYDFIMEPPKNFNTLVGEKG 460  
DB 469 ilgvsgqetvalvngscgkstvtqlmqrllydptegmvsvdgqdditlrnfrlre 528

OY 461 AQMSGGOKRIATARALVRNPKILLIDEATSALDSESKSAVQAALAK 507  
 Db 529 aqisggqkqriataralvrnpkillideatsaldeseavqaalck 575

## RESULT 10

AA78879  
 ID AAB78879 standard; Protein; 1275 AA.

AA78879;

19-MAY-2000 (first entry)

Rat mdrlb2 (multispecific drug transporter) protein sequence.

Rat: mdrlb2; multispecific drug transporter; cancer; inflammation;  
 cardiovascular disease; central nervous system disorder;  
 autoimmune disease; kidney disease; drug formulation.

Rattus sp.

US6025160-A.

15-FEB-2000.

22-JUL-1998; 98US-0120513.

22-JUL-1998; 98US-0120513.

(SMIK ) SMITHKLINE BEECHAM CORP.

Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;

WPI: 2000-181810/16.

N-PSDB; AA290198.

Isolated polynucleotide encoding a rat mdrlb2 drug transporter  
 polypeptide, useful for treatment of e.g. cancer, autoimmune disease,  
 central nervous system disorders

Claim 6; Column 17-20; 17pp; English.

This sequence represents a rat mdrlb2 multispecific drug transporter  
 protein. The mdrlb2 nucleotide sequence has cytosolic, antiinflammatory,  
 cardiac, neuroprotective, immunosuppressive and nephrotropic activity.  
 Understanding the functioning of the mdrlb2 polynucleotide and protein in  
 transgenic animal models is useful for treating and preventing diseases  
 such as cancer, inflammation, cardiovascular disease, central nervous  
 system disorders, autoimmune disease, and kidney disease. The use of the  
 protein in cell based, membrane based, or binding assays may enhance drug  
 formulation, selection of formulation excipients and compound design.

Sequence 1275 AA;

Query Match 60.8%; Score 1576; DB 21; Length 1275;  
 Best Local Similarity 58.4%; Pred. No. 1.7e-137;

Matches 308; Conservative 92; Mismatches 105; Indels 22; Gaps 4;

OY 1 MITGLASLVNGACLPMLPLVLGEMSDNL-----SGCLVQNTNTYFFRL-- 45

Db 49 malgtlaalngltlplmlvfygmdstftgaetrllpsvngselnstgvsdsleed 108

OY 46 -----TLVYVGIVAAITFGVYIOISLWITTAARQTRIKQFHSVLAQDGMPSQDICE 101

Db 109 manyayyvtgagvllvayvswlaagrqhkrqkflthalmqelwtdvndage 168

OY 102 LNTNMT-DIDKISDGIQDKIALLFQNMSTFSGIAGLVKGMKLTLYLSTSPILMASAA 160

Db 169 lntltldvsktndgigdklqgmffgstltftagflifglsqwkltlvllavspilglsaa 228

OY 161 ACSRMVLSLTSKELSAYSKAGAAVEVTSRTVTYAFPAQOEKLORYQNTLKDADPSIK 220

Db 229 mnavltstfnkelqayakagavaeevlaairvialfggqkkelerynknlceakrvglk 288  
 OY 221 RTIASKVSLGAAYFFMNGSYGLAFWYGTSLILNGEGYITGYLAVFVSHSTICGA 280  
 Db 289 kaitanlsiglayllvyasafafwytstivlsne--ysigvltvfsslllgtfslghl 346  
 OY 281 VPHFEMFARGAAFHIFOVYIDKKPSIDNFSSTGKVECTEVEERKNVSFNVPSPSIR 340  
 Db 347 apnieafanagaayelfkldnepsidstfkghpdsimglftknylfnpsrsevk 406  
 OY 341 ILKGLNLRKISGEFVALVGLNGSGKSTVVQLLQRLYDPDDGFIMVDENDIRALNVRHYRD 400  
 Db 407 ilkglnlkvsgqvalvngsgcgskettvqllyrlydpdlegvslgqdlrtlnvryle 466  
 OY 401 HIGVSGEPVLFQTTISNNIKRYGRDVTDEMEPAAREANAYDFMEFPNKFTLYGEG 460  
 Db 467 llygvsgepllfatlaenlrygrenvmdetkavkeanaydfimlpkftnlvgerg 526  
 OY 461 AQMSGGOKRIATARALVRNPKILLIDEATSALDSESKSAVQAALAK 507  
 Db 527 aqisggqkqriataralvrnpkillideatsaldeseavqaalck 573

## RESULT 11

AAB60409  
 ID AAB60409 standard; Protein; 1275 AA.

AAB60409;

24-APR-2001 (first entry)

Rat mdrlb2 multidrug resistance protein, SEQ ID NO:2.

Rat: multidrug resistance; mdrlb2; multi-drug transporter family;  
 drug screening; pharmacokinetic analysis; oral absorption;  
 formulation design; bioavailability; transgenic animal; knockout animal;  
 inflammation; cardiovascular disease; central nervous system disorder;  
 cancer; autoimmune disorder; kidney disease.

Rattus sp.

US6169166-B1.

02-JAN-2001.

29-NOV-1999; 99US-0450105.

22-JUL-1998; 98US-0120513.

(SMIK ) SMITHKLINE BEECHAM CORP.

(SMIK ) SMITHKLINE BEECHAM PLC.

Chenery RJ, Ellens H, Brun KA, Yue L, Feild JA;

WPI: 2001-158280/16.

N-PSDB; AAF27498.

New rat multidrug resistance protein (mdrlb2), useful for screening or  
 identifying compounds that are (ant)agonists/inhibitors of the  
 ratmdrlb2, as well as compounds with optimal development  
 characteristics

Claim 1; Column 17-24; 13pp; English.

The invention relates to a novel rat multidrug resistance (mdr)  
 protein, mdrlb2 (AAB60409). Multi-specific drug transporter family  
 proteins are present in cell which have a barrier function, such as  
 intestinal epithelial cells, brain microvessel endothelial cells,  
 kidney epithelial cells, and liver hepatocytes, and are also expressed  
 by certain cancer cells. The rat mdrlb2 protein is useful for  
 screening or identifying compounds that are agonists or antagonists  
 of mdrlb2 activity. It may also be used to establish assays to  
 predict oral absorption and pharmacokinetics of drugs in humans, and



CC thus enhance the design of formulations through the identification of  
 CC compounds with optimal development characteristics (i.e., high oral  
 CC bioavailability, UID (once a day) dosing, reduced drug interactions,  
 CC reduced variability, and reduced food effects), specifically to  
 CC avoid interactions with human mdr-1. Transgenic and knockout animals  
 CC created interactions encoding the rat mdr1b2 may be used to gain an  
 CC insight into treating and preventing human diseases such as cancer,  
 CC inflammation, cardiovascular disease, central nervous system disorders,  
 CC autoimmune disorders and kidney disease. The present sequence represents  
 CC rat mdr1b2.

XX Sequence 1275 AA;

Query Match 60.8%; Score 1576; DB 22; Length 1275;  
 Best Local Similarity 58.4%; Pred. No. 1.7e-137;  
 Matches 308; Conservative 92; Mismatches 105; Indels 22; Gaps 4;

OY 1 MILGILASLVNGACPLMLPVLGEMSDNLI-----SGCLVOTNTSFRFL--- 45  
 DB 49 malgltaaihhgtlpllmvfygymtdstfgaetrlpsvtngselnstqvsdsleed 108  
 OY 46 ----TLVYVGIVAAALFEGYIOISLWITTAAROTKRIRKQFHFHSLAODIGWFDSCDICE 101  
 DB 109 mamaygyyvgiagvliavaylqvsfawcslaagrgyihkirkqffhalmngelgwfvdvndage 168  
 OY 102 LNTFRMT-DIDKISDGIKIDKIALLFQNMSTFSGIAGLVGKMKLTVLTSTSPILMASAA 160  
 DB 169 lntfrltdvskindgldlqmfifqstltsagfllgflsgvklvtlilavspilglssa 228  
 OY 161 ACSRWISLTSKELSAVSKAGAAVEVLSSIRTVIAFRQKELEORYTONLKDARFGIK 220  
 DB 229 mwakvlstfnkelqayakagavaeevlaaitvtaifgqkkelerynkhleakrvyglk 288  
 OY 221 RTIASKVSLGAVYFFMNGTYGLAFWYGTSLINGEPGYTIGTVLAVFESVHSSYCIGAA 280  
 DB 289 kaitanisigayllvyasaayalfwygtslvsne--ysigqylvltvfisllgltfsgl 346  
 OY 281 VPHFEFAIARGAAFHIFQVIDKPSIDNFSIAGYKPESEIGTEVERKNVSPNPSRSIK 340  
 DB 347 apulaefanaragaayelfiklfnepsidsfstckghkpsidmgnlftknyfhyprsevk 406  
 OY 341 IKGLNLRIKSGEYVALVGLNGSGSTVQLOLRLYDPDPPGFLMVENDIRALNVHRYND 400  
 DB 407 ltkglnlkvsgqvalvaynsgcgkstvtvqlqrlpdyplegevsldgqdlrtlnvlyle 466  
 OY 401 HIGVVSQEPVLFGTTSNNIKYGRDVTDEMERARAREANAYDFIMEFNKFTLVGEGK 460  
 DB 467 llygvvsqepvlftatlaenlyrgrenvmdeldekavkenaydfimklphkftltvgerg 526  
 OY 461 AQMSGGOKRIARALVNRPKILLIDETSALDSEKSAVAQALEK 507  
 DB 527 aqlsggqkgrlataralvnrpkllllldeatsaldtesaavvaaidk 573

RESULT 12  
 AAM48997 standard: Protein, 1280 AA.

AC AAM48997;  
 DT 23-SEP-1998 (first entry)

DE Wild-type human P glycoprotein (Pgp).

KM Human P glycoprotein; Pgp; multi-drug resistance; cancer;  
 KW UIC2 monoclonal antibody; mAb; cytotoxic; transmembrane efflux pump.

OS Homo sapiens.  
 XX WO9821325-A1.  
 XX 22-MAY-1998.

XX 17-NOV-1997; 97WO-US21214.  
 PF 15-NOV-1996; 96US-0752447.  
 PR (INGE-) INGENEX INC.  
 PA (UNIT ) UNIV ILLINOIS FOUND.  
 XX Mechner E, Roninson IB;  
 XX WPI, 1998-297930/26.  
 DR N-PSDB; AAV32645.  
 PT Immunological reagent specific for P-glycoprotein - useful for  
 PT detecting multi-drug resistant cancer, isolating hematopoietic  
 PT cells and selective cell killing  
 PS Disclosure: Pages 38-42; 89pp; English.

The present sequence represents the wild-type human P glycoprotein (Pgp) protein. Pgp is a transmembrane efflux pump protein involved in multi-drug resistance of cancer cells. The invention provides methods for developing and using immunological reagents specific for certain mutant forms of Pgp and wild-type Pgp in a conformation associated with substrate binding or in the presence of ATP depleting agents. An example of the immunological reagent is the UIC2 monoclonal antibody (mAb). mAb UIC2 specifically binds to Pgp in a particular biochemical conformation and is capable of inhibiting drug efflux from Pgp-expressing cells. The immunological reagents are claimed to be useful for detecting Pgp expression in mammalian cells, including low level expression, particularly in cancer cells to diagnose multi-drug resistance. The invention claims that these immunological reagents are more specific than known reagents for detecting Pgp and they also eliminate the need for costly and laborious screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.

SO Sequence 1280 AA;

Query Match 60.8%; Score 1576; DB 19; Length 1280;  
 Best Local Similarity 57.9%; Pred. No. 1.8e-137;  
 Matches 305; Conservative 101; Mismatches 99; Indels 22; Gaps 5;

OY 1 MILGILASLVNGACPLMLPVLGEMSD-----NLISGLVOT--NTYSRF----- 43  
 DB 51 mvvgltlaaihhgaglpilmvfygemldifanagnledlnstnrxdsldtqfimmleed 110  
 OY 44 --RLTVYVGIVAAALFEGYIOISLWITTAAROTKRIRKQFHFHSLAODIGWFDSCDICE 101  
 DB 111 mctayayysglaagvliavaylqvsfawcslaagrgyihkirkqffhalmngelgwfvdvndage 170  
 OY 102 LNTFRMT-DIDKISDGIKIDKIALLFQNMSTFSGIAGLVGKMKLTVLTSTSPILMASAA 160  
 DB 171 lntfrltdvskinegldlqmfifqstltsagfllgflsgvklvtlilavspilglssa 230  
 OY 161 ACSRWISLTSKELSAVSKAGAAVEVLSSIRTVIAFRQKELEORYTONLKDARFGIK 220  
 DB 231 mvakllstfdkellayakagavaeevlaaitvtaifgqkkelerynkhleakrvyglk 290  
 OY 221 RTIASKVSLGAVYFFMNGTYGLAFWYGTSLINGEPGYTIGTVLAVFESVHSSYCIGAA 280  
 DB 291 kaitanisigaeallvyasaayalfwygtslvsne--ysigqylvltvfisllgltfsgl 348  
 OY 281 VPHFEFAIARGAAFHIFQVIDKPSIDNFSIAGYKPESEIGTEVERKNVSPNPSRSIK 340  
 DB 349 spulaefanaragaayelfiklfnepsidsfstckghkpsidmgnlftknyfhyprsevk 408  
 OY 341 IKGLNLRIKSGEYVALVGLNGSGSTVQLOLRLYDPDPPGFLMVENDIRALNVHRYND 400  
 DB 409 ltkglnlkvsgqvalvaynsgcgkstvtvqlqrlpdyplegevsldgqdlrtlnvlyle 468  
 OY 401 HIGVVSQEPVLFGTTSNNIKYGRDVTDEMERARAREANAYDFIMEFNKFTLVGEGK 460





[illegible]

```
AC AAE00304;
XX
XX 13-JUN-2001 (first entry)
DT
XX
DE Dog P-glycoprotein (PGP) #2.
XX
KW Dog: P-glycoprotein; PGP; multidrug transporter; MDRL;
XX drug bioavailability; transgenic animal; genetic model.
XX
OS Canis familiaris.
XX
PN W0200123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000OWO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
DR WPI: 2001-235373/24.
XX N-PSDB; AAD03489.
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
XX Example 2; Page 72-75; 11pp; English.
PS
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDRL) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. Pgp and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC
CC The present sequence is dog P-glycoprotein (PGP). The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 1280 AA;

Query Match          60.8%; Score 1576; DB 22; Length 1280;
Best Local Similarity 57.1%, Pred. No. 1,8e-137;
Matches 302; Conservative 98; Mismatches 105; Indels 24; Gaps

    1 MITGLIASLVNACPLMPVLVGENSDNLISCL-----VQNTYSFRR--- 44
      ||| :|||:||||| | | | | | | | | | | | | | | | | | | |
Db 50 mlvgsmaaiibgaaplmlmlyfgmntdstandgisnrtkfpvlinessinnqfinhle 109
      ||| | | | | | | | | | | | | | | | | | | | | | | | |
    45 -----LTLVYVGIGVAALIEGTIQLSLWITTAAROTKRIKOPFNHSLVLAODIGMPEDSCI 99
      || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110 eemtyayayysgiagvivaayiygvsfwclaargllkkirkbfinaimgexfivfhav 169
      || | | | | | | | | | | | | | | | | | | | | | | | | |
    100 GELNTRMT-DIKISGIDGCTALLFFONNSTFSIGAVLVGMVKLTLTLSPLIMAS 158
      |||||:| |:|:|||||:| |:|:| | | | | | | | | | | | | |
Db 170 ginctlctddskinegydkvgmfifgsiatftfcgfvyafcpwkltlvialispligls 229
      |||||:| |:|:|||||:| |:|:| | | | | | | | | | | | | |
    159 AAGCRMVISTLSKELSAISKAGAFAEEVLIIRTVIAFRAGEKELORYTONLKDAKDFG 218
      || :|| :| || | | | | | | | | | | | | | | | | | | | |
Db 230 aaiwlkissfdckellayekagavaeeelaairtviafatgqgkclerynkleeakrig 289
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:35:53 ; Search time 42.7 Seconds  
(without alignments)  
916.949 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592

Sequence: 1 MUGILASLVNGACLEPLMP.....SESKSAVQAALKEKTPRYSF 514

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1590	61.3	1276	1	DVMS1 multidrug resistan
2	1579	60.9	1280	1	DVHU1 multidrug resistan
3	1571.5	60.6	1277	2	JH0502 P-glycoprotein - r
4	1565.5	60.4	1276	2	A34786 multidrug resistan
5	1564	60.3	1276	1	DVHY1C multidrug resistan
6	1546.5	59.7	1287	2	S55692 multidrug resistan
7	1479	57.1	1279	1	DVHU3 multidrug resistan
8	1464	56.3	1281	2	148123 multidrug resistan
9	1463	56.4	1276	1	DVMS2 multidrug resistan
10	1445	55.7	1278	2	S41646 multidrug resistan
11	1301	50.2	1104	1	DVMS1A P-glycoprotein - r
12	1243.5	48.0	1321	2	T42228 P-glycoprotein resist
13	1243.5	48.0	1321	2	T42842 P-glycoprotein resist
14	1158	44.5	1294	2	T19982 multidrug resistan
15	1050.5	40.7	1275	2	T31073 multidrug resistan
16	1025.5	39.6	1302	2	A41249 multidrug resistan
17	1023	39.5	1321	2	T23476 multidrug resistan
18	1021	39.4	1321	2	S27337 multidrug resistan
19	1017	39.2	1286	2	T02187 multidrug resistan
20	1008	38.9	1230	2	E85023 probable ABC trans
21	1000.5	38.6	1283	2	A47377 P-glycoprotein resist
22	998	38.5	1292	2	T48007 P-glycoprotein resist
23	996	38.4	1229	2	D85023 P-glycoprotein resist
24	993	38.3	1229	2	T52319 P-glycoprotein resist
25	985.5	38.0	1222	2	T14805 P-glycoprotein resist
26	985.5	38.0	1222	2	T14805 P-glycoprotein resist
27	985	38.0	1323	2	H85202 P-glycoprotein resist
28	977.5	37.7	1318	2	T21266 P-glycoprotein resist
29	975.5	37.6	1408	2	T43261 P-glycoprotein resist

30	972.5	37.5	1254	2	S27338 P-glycoprotein C -
31	972.5	37.5	1268	2	T22094 hypothetical prote
32	971.5	37.5	1266	2	T22094 hypothetical prote
33	966	37.3	1229	2	F68155 probable ABC trans
34	965.5	37.2	1286	2	A42150 P-glycoprotein pgp
35	950.5	36.7	1302	2	S30327 P-glycoprotein resist
36	945.5	36.5	1362	2	T41534 P-glycoprotein resist
37	944.5	36.4	1310	2	S30328 P-glycoprotein resist
38	940.5	36.3	1245	2	G86404 probable P-glycopr
39	928	35.8	1341	2	JG0166 LAMP1 protein - L
40	923.5	35.6	1158	2	T25082 hypothetical prote
41	923	35.6	1302	2	B41249 multidrug resistan
42	920	35.5	1307	2	T30882 multidrug resistan
43	916.5	35.4	1270	2	T21269 hypothetical prote
44	915.5	35.3	1327	2	T21268 hypothetical prote
45	913.5	35.2	1233	2	T04251 P-glycoprotein 2 -

## ALIGNMENTS

RESULT 1  
DVMS1  
multidrug resistance protein 1 - mouse  
N:Alternate names: P-glycoprotein 1  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 19-Jan-2001  
C:Accession: A33719; A25057; I57510  
R:Raymond, M.; Gros, P.  
Proc. Natl. Acad. Sci. U.S.A. 86, 6488-6492, 1989  
A:Title: Mammalian multidrug-resistance gene: correlation of exon organization with s  
A:Reference number: A33719; MUID:89367274  
A:Accession: A33719  
A:Molecule type: DNA  
A:Residues: 1-1276 <RAY>  
R:Gros, P.; Croop, J.; Housman, D.  
Cell 47, 371-380, 1986  
A:Title: Mammalian multidrug resistance gene: complete cDNA sequence indicates strong  
A:Reference number: A25057; MUID:87028229  
A:Accession: A25057  
A:Molecule type: mRNA  
A:Residues: 1-1276 <GRO>  
R:Cross-references: GB:M44757; NID:q199100; PIDN:AA79005.1; PID:q387426  
A:Accession: I57510  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-21 <RES>  
A:Cross-references: GB:M60348; NID:q199102; PIDN:AA39513.1; PID:q554199  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
structurally and functionally unrelated lipophilic antitumor drugs.  
C:Genetics:  
A:Gene: mdr1 (pgp1)  
A:Introns: 21/2; 38/3; 92/1; 111/2; 175/2; 233/3; 274/2; 332/3; 370/3; 407/3; 449/3;  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane  
F:1-637; 653-1276/Region: duplication  
F:409-603/Domain: ATP-binding cassette homology <ABC1>  
F:426-433/Region: nucleotide-binding motif A (P-loop)  
F:550-554/Region: nucleotide-binding motif B  
F:1051-1247/Domain: ATP-binding cassette homology <ABC2>  
F:1068-1075/Region: nucleotide-binding motif A (P-loop)  
F:1194-1198/Region: nucleotide-binding motif B  
F:73; 91; 96; 103/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:432/Binding site: ATP (Lys) #status predicted  
F:1074/Binding site: ATP (Lys) #status predicted

Query Match 61.3%; Score 1590; DB 1; Length 1276;  
Best Local Similarity 57.7%; Pred. No. 2.5e-104;

Matches 307; Conservative 99; Mismatches 94; Indels 32; Gaps 5;

```
Oy 1 MILGLIASLVNGACPLMPVLVLEKMSDNL-----SGCLVQNT----- 39
Db 50 MILCTLAIIHGTLPLMLVFGNMSTFTKAASILPSITNOSGPNSTLIISNSSLEE 109
Oy 40 ---SSEFRLTYLVYGVAALIFGYIOISLWITAAAROTKRIKROFPSPVAADIGMFD 96
Db 110 MAITVA-----YYGIGGCVLITVAITYOVSIMWCLAAGROIHKIKRFPALNOELGMFDV 164
Oy 97 CDIGELNTRMT-DIDIKSIDGIDKIALIFQNMSTFSGISGLAVGLKGMKTLVLTSTPLI 155
Db 165 HDVGLNTRLRLDDVSKINDGIDGKIGMFQSIITFFLAGFIIGFISGMKTLVILAVSPLI 224
Oy 156 MASAACSRMVIISLTSKELSAVSKAGAAVEVLSIRTVIAFRAOEKELQRYTONLDAK 215
Db 225 GLSSALMAKVLTSFTNKLQAYAKAGAAVEVLAIRTVIAFGQOQLEERYNKNLEBAK 284
Oy 216 DFGIKRTASKVSIGAVYFFPMNGTYIGLAFWYGTSTLIILNGEGYITIGYLAFFESYIHSY 275
Db 285 NVGIKKATTAISISIGIALLVAYASALAFWYGTSTLVLSNE--YSIGELTVTFEFLILGTF 342
Oy 276 CIGAAVPEFEFFAIFARAAAFHIFQYIDKKPSINDPSTAGGYPESEEGVEKKNVSPNPS 335
Db 343 SIGHLAPITEFAFNANKGAFFELFKIIDNPSIDSTSTKGYKPDSTMGLEKKNVHFNPS 402
Oy 336 RPSIKILKGLMLRIKSGEVTVALVGLNGSGKSTVVOQLQRLDPPDGFMTVDENDIRALNV 395
Db 403 RSEVQILKGLMLKVKSGQVTVALVNGSGCKSTYQMLQRLDPLEGVYSIDGQDIRITNV 462
Oy 396 RHYRDHIGVSOEPEVLFETGTTISNNIKYGRDVTDEEMRAAREANADFIEMFENKENTL 455
Db 463 RYLRLEIGVSOEPEVLFATTTIAENIRGREDEVTDLEKAKVENANADFIKLPHOEDTL 522
Oy 456 VGEKAQMSGOKORIIATARALVRNPKILLIDDEATSDLSKSSAVOALAEK 507
Db 523 VGERAQISGCKOKRIATARALVRNPKILLIDDEATSDLTSEAVVOALAEK 574
```

RESULT 2

DVHUT

multidrug resistance protein 1 - human

N:Alternate names: P-glycoprotein 1

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence\_revision 18-Aug-1995 #text\_change 19-Jan-2001

C:Accession: A34914; PS0162; S15500; A25059; S43838; I52238; I65204

R:Chen, C.; Clark, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.

J. Biol. Chem. 265, 506-514, 1990

A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin

A:Reference number: A34914; MUID:90094448

A:Accession: A34914

A:Molecule type: DNA

A:Residues: 1-1280 <CHE>

A:Cross-references: GB:M29447; GB:J05168; NID:q187496; PIDN:AAA59576.1; PID:q386862

R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.

submitted to JIPID, April 1991

A:Reference number: PS0162

A:Accession: PS0162

A:Molecule type: DNA

A:Residues: 1-22 <KIO>

R:Kloka, N.; Yamano, Y.; Komano, T.; Ueda, K.

submitted to the EMBL Data Library, April 1991

A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) expression

A:Reference number: S15500

A:Accession: S15500

A:Molecule type: DNA

A:Residues: 1-22, 'R' <K12>

A:Cross-references: EMBL:X58723; NID:q34522; PIDN:CAA4158.1; PID:q34523

R:Chen, C.; Chin, J.E.; Ueda, K.; Clark, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I.

Cell 47, 381-389, 1986

A:Title: Internal duplication and homology with bacterial transport proteins in the mdrl

A:Reference number: A25059; MUID:87028230

A:Accession: A25059

```

A:Molecule type: mRNA
A:Residues: 1-184, 'V', 186-1280 <CH2>
A:Cross-references: GB:M4758; NID:g187468; PIDN:AAA59575.1; PID:g307180
R:Chambers, T.C.; Pohl, J.; Glass, D.B.; Kuo, J.F.
Biochem. J. 299, 309-315, 1994
A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase
A:Reference number: 54838; MUID:94220047
A:Accession: 54838
A:Molecule type: protein
A:Residues: 656-689 <CHA>
R:Geckeler, V.; Meger, S.; Probst, H.
Biochem. Biophys. Res. Commun. 169, 796-802, 1990
A:Title: mdr1/P-glycoprotein gene segments analyzed from various human leukemic cell
A:Reference number: 152238; MUID:90290529
A:Accession: 152238
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 178-215 <RES>
A:Accession: 165204
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 800-856 <RE2>
A:Cross-references: GB:M37725; NID:g183537; PIDN:AAA88047.1; PID:g553315
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant C
structurally and functionally unrelated lipophilic antitumor drugs.
C:Genetics:
A:Gene: GDB:PCY1; MDR1
A:Cross-references: GDB:120712; OMIM:171050
A:Map position: 7q21-7q21
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein
F:1-638, 653-1280/Region: duplication
F:49-350/Domain: hydrophobic <HB1>
F:351-637/Domain: hydrophilic <HL1>
F:410-604/Domain: ATP-binding cassette homology <ABC1>
F:427-434/Region: nucleotide-binding motif A (P-loop)
F:531-535/Region: nucleotide-binding motif B
F:638-708/Domain: linker <LIN>
F:709-993/Domain: hydrophobic <HB2>
F:994-1280/Domain: hydrophilic <HL2>
F:1053-1249/Domain: ATP-binding cassette homology <ABC2>
F:1070-1077/Region: nucleotide-binding motif A (P-loop)
F:1136-1200/Region: nucleotide-binding motif B
F:91, 94, 99/Binding site: carboxylate (Asn) (covalent) #status predicted
F:433/Binding site: ATP (Lys) #status predicted
F:667, 667, 671/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status
F:667, 671, 683/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #st
F:1076/Binding site: ATP (Lys) #status predicted

Query Match 60.9%; Score 1579; DB 1; Length 1280;
Best Local Similarity 58.1%; Pred. No. 1,56-103;
Matches 306; Conservative 100; Mismatches 99; Indels 22; Gaps 5;

1 MILGLIASLVNGACPLPLPLVLEMSD-----NLISCLVQT--NTYSFF----- 43
QY :|||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 11
Db 51 MYYGTLAIHTIGAGPLPLMLVLFGEMTIDFANVNGNLEDSNITKRSYSDINDGPFMNLLED 110
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 --RLTLVVGIGVAAALTCGYTIQISLMTTAAQRKRIKQFFHSYLAODIWFPSQDGE 101
QY :|||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 11
Db 111 MTRVYVYSGIGAGVLAAYIYOVFWCMLAARQIHRIKQFFHAIMROELGMPVHDGE 170
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
102 LNTFRMT-DIDKISDQIGKIALPQNMSTFSGIAGVGVKCMKTLVLVLTSPILMASAA 160
QY :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 11
Db 171 LNTFRITDVSKNBSIGKIGTFQSMATFTFGFLVGTGRKMKTLVLVLTSPVLGSSAA 230
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 ACSRAVISLTKNELSAVSKAGAVAEVLSSIRTVIAFRAQEKELQRYTQNLKADKFGIK 220
QY :||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 11
Db 221 VWAKLTLSFTDKELLAVYKAGAVAEVLAAIRTVIAFGGQKKELEERYKNLEAKRIGIK 290
:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 11
221 RTIASKVGLGAVYFMNGTIGLAIFYGTSILNGEPGTYIGTVLAFFSVTHSSYICGAA 280
:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 11

```

Db 291 KATPANSIGAAFLILTVASVALAFWYGTLLVSGE - YSIOQVLVFEFSVLIGAFSVGQA 348

Qy 281 VPHRETFPIAAGAAPHIEQVLDKPSIDNESTAGCKPESIGCYEEFKVNSFNYSRPSIK 340

Db 349 SPSEIAFANAGAAEIEKIIDNKPISIDYSKSGSKGHPNIGNLEFRNVHFSYPSRKEVK 408

Qy 341 ILKGNILNKSGEFTALVGLNGSGKSTVYVOLLQRYDDDDQFIWDEDIDALANVRHFD 400

Db 409 ILKGNILNVGSGQVVALVGLNGSGCKSTVYVOLLQRYDDDEAMVSVDGDIPTINRFURE 468

Qy 401 HIGVVSQEPVLEFGTTISNNIKYGRBDYDDEMERARAENAYDFMEPNKFNFTLVGSKG 460

Db 469 IIGVVSQEPVLEFATTIAINIKYGRBNVTMDLEKRAVKANAYDFIMKPLPHKFDTLVGERG 528

Qy 461 AOMSGGOKORIAIARALVNRPKILLDDEATSAALDESKSAVOALEK 507

Db 529 AOLSGGOKORIAIARALVNRPKILLDDEATSAALDESEAVVOVALDK 575

RESULT 3

JH0502

P-glycoprotein - rat

N:Alternate names: multidrug resistance protein mdr1b

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 02-Feb-2001

C:Accession: JH0502; S22353; S22352

R:Silverman, J.A.; Raunio, H.; Gant, T.W.; Thorgerirsson, S.S.

A:Title: Cloning and characterization of a member of the rat multidrug resistance (mdr)

A:Reference number: JH0502; MUID:9203081

A:Accession: JH0502

A:Molecule type: mRNA

A:Residues: 1-1277 <SIL>

A:Cross-references: GB:M62425

R:Deuchars, K.L.; Dutchie, M.; Ling, V.

Biochem. Biophys. Acta 1130: 157-165 1992

A:Title: Identification of distinct P-glycoprotein gene sequences in rat.

A:Reference number: S22351; MUID:92223089

A:Accession: S22353

A:Molecule type: DNA

A:Residues: 1212-1226, '1', 1228-1277 <DEU>

A:Cross-references: EMBL:X61104; NID:956890; PIDN:CAA43416.1; PID:g1334219

A:Accession: S22352

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1212-1226, '1', 1228-1270, 'SV', 1271-1277 <DEZ>

A:Cross-references: EMBL:X61103; NID:956888; PIDN:CAA43415.1; PID:g1334218

C:Comment: This protein has the nucleotide binding motifs and ATP binding active transpo

C:Genetics:

A:Gene: mdr1b

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-loop

F:409-603/Domain: ATP-binding cassette homology <ABC1>

F:432-474/Region: nucleotide-binding motif A (P-loop)

F:550-554/Region: nucleotide-binding motif B

F:1053-1248/Domain: ATP-binding cassette homology <ABC2>

F:1070-1078/Region: nucleotide-binding motif A (P-loop)

F:1195-1199/Region: nucleotide-binding motif B

F:433/Binding site: ATP (Lys) #status predicted

F:1076/Binding site: ATP (Lys) #status predicted

Query Match 60.6%; Score 1571.5; DB 2; Length 1277;

Best Local Similarity 58.1%; Pred. No. 5.2e-103;

Matches 307; Conservative 94; Mismatches 104; Indels 23; Gaps 4;

Qy 1 MILGILASLVNACGLPLMPVLGEMSDNLI-----SGCLVQNTYSEFR- 45

Db 49 MALGTLAAIHHTPLPLMLTVGVMYTDSEFSPDRDPSHSDRAITNQSEINSTHTVSDTLEE 108

Qy 46 -----TLVYVGICVAALLFGYTIQISLMTITAAQRKRRKQFHSVLAODIGMFDSCDIG 100

Db 109 DMAAYAYVYTGAGVAILVAIYGVSLWCLAGROVRIHKIRKQFHAIMNQGEMFVNDAG 168

Oy 101 ELNMT-DIDKISGIDCKTALLFQNMSTPESIGLAVGLVGMKLTIVTISTPLIMASA 159  
| : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 169 LFNRLFDVDVKINDIGDIDKLGMEFFQSITTFISAGFIIFIGMKLTIYILAVSFLIGLS 228  
Oy 160 AACSRMVTSLSKELISAASKAGAAVEEVLSSIRTVIAFRAOEKELORTONLKADKDFGI 219  
| : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 229 AMMAKVLIISFNKEQLQAAYAKAAGAAVEEVLAARIVYIARGCGKELEERYNNKLAEAKRGI 288  
Oy 220 KRTIASKVSICGVFPFMNGTYGLAFWYGTSLILNGEPCTTGTYLANAFVSIIHSYCIGA 279  
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 289 KKAIANISIGIAYLYLVASYALAFWYGTSLVLSNE--YSYGVOYLTVFFSIILCTFSIGH 346  
Oy 280 AVPHETFAIRAAGAFTHEFYVIDKKPSIDNFNSTGYKRRESIEGVPEEFNVFNSPRPSI 339  
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 347 LAPNFAAPNARGAAVELFKITIIDNEPDSIDSTYGHKDJSMGNLFENNYFNPSPSEV 406  
Oy 340 KILKGLNRIRKSGETVALVGLNGSGKSTVVOLLORLYDDPDGFTMVENDIRALNVRHYR 399  
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 407 KILKGLNLKVKVSGQVALVGLNGSGCKSTYYOLLORLYDPICEEVSIDGDITINVRYL 466  
Oy 400 DHICVSEPVLPFGTTISNNIKRYGRDYTDDEMERAAEAANAYPIMEFPKPFMTLVGEK 459  
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 467 EIIGVSEPEVLPFAATTIAENIRYGRENTMDIEKAIVEANAAYDFIMKLPRKPTLVGER 526  
Oy 460 GAOMSGGOKORIAARALVNRPKIITIIDATSALDSEKSASVAOLEK 507  
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 527 GAOLSGGOKORIAARALVNRPKIITIIDATSALDTEBEAVOALDK 574

RESULT 4  
A34786  
multidrug resistance protein 1a - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 02-Feb-2001  
C:Accession: A34786; A35671  
R:devaault, A.; Gros, P.  
MOL. Cell. Biol. 10, 1652-1663, 1990  
A>Title: Two members of the mouse mdr gene family confer multidrug resistance with ov  
A:Reference number: A34786; MUID:90205845  
A:Accession: A34786  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1276 <DEV>  
A:Cross-references: GB:M30697; NID:g19911; PIDN:AAA39517.1; PID:g387429  
R:Hsu, S.T.H.; Cohen, D.; Kirschner, L.S.; Lotheisen, L.; Hartstein, M.; Horwitz, S.B.  
MOL. Cell. Biol. 10, 3596-3606, 1990  
A>Title: Structural analysis of the mouse mdra (P-glycoprotein) promoter reveals the  
A:Reference number: A35671; MUID:90287150  
A:Accession: A35671  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-525, 'HV', 528-938, 'S', 940-1035, 'V', 1037-1276 <HSHU>  
A:Cross-references: GB:M33581; NID:g1919104; PIDN:AAA39514.1; PID:g387427  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; P-1  
F:408-600/Domain: ATP-binding cassette homology <ABC1>  
F:423-431/Region: nucleotide-binding motif A (P-loop)  
F:547-551/Region: nucleotide-binding motif B  
F:1049-1245/Domain: ATP-binding cassette homology <ABC2>  
F:1066-1074/Region: nucleotide-binding motif A (P-loop)  
F:1192-1196/Region: nucleotide-binding motif B  
F:429/Binding site: ATP (Lys) #status predicted  
F:1072/Binding site: ATP (Lys) #status predicted

Query Match 60.4%; Score 1565.5; DB 2; Length 1276;  
Best Local Similarity 56.5%; Pred. No. 1,4e-102;  
Matches 299; Conservative 103; Mismatches 98; Indels 29; Gaps 4;

Oy 1 MILGLIALVNCACPLMPVLYGEMSDNISGLCYQT-----T 39  
| : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :  
Db 50 MLVGVIATLHHVALPMLMLIFGDMTDSFASGVNKSNTNMSEADKRAMFAKTEEMTT 109

```
OY 40 YSEFRLTYVYGIVAAALIFGYIOISLMTITPAROTKRIRKOFPHSVLAODIGFSDCI 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 YAY-----YTGIGAGVLIVATVGSFMCCLAAGROIHKIRQFHAIMNOEIGMFDVHDV 164
OY 100 GELNTRMT-DIDKISDIDGDKIALLFONMSFFSIGLAAGLVGKMLTLVLTSTSPILMAS 158
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 165 GELNTRLTDVDSKINEGIDGKIGMFQAMATPFEGFIIIGFTRGKMLTLVILAIISPVLGIS 224
OY 159 AACGRMVISLTSKLSATSKAGAAVEVLSIRTVIAFRAOEKLOXYTQMLKAKDFG 218
    | : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 225 AGIWKILISSEFDKELHAYAKAGAAVEVLAIRTVIAFGGOKKELERNNNLEBAKRLG 284
OY 219 IKRTASKVSLGAVYFFMNGTGYLAFWYGTSLIINGEPGYTIGTVLAFFSVHSYCSYC 278
    || : : || : : || : : || : : || : : || : : || : : || : : || : : || : :
Db 265 IKKATINISKAAPALLIYASVALAFWYGTSLVYSKE--YSTGVLYVFFSVLTAFSVG 342
OY 279 AAVPFETFAIARGAAPHIPOVIDKKPSIDNESTAGYKPESTIEGTEVERKNVSFNPSRPS 338
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 343 QASPMIEAFANARGAAYEFKLIIDNKPSIDSFSGKHKPDNIOGMLERKNHFSYPSRKE 402
OY 339 IKILKGLNLRKISGETVALVGNLNGSGKSTVQQLRLDLPDGFIMVDENDIRALNVRYL 398
    : : ||||| : : ||||| : : ||||| : : ||||| : : ||||| : : ||||| : : |||||
Db 403 VOILKGLNLRKYSQGVVALVGNLNGSGKSTVQQLMQLRPLDGMVSDGQDIRTNVRYL 462
OY 399 RDHIVGVSQEPVLFGTTSNNIKRYGRDVTDEMERARAEANAYDFIMEFPKFNVLVGE 458
    | : ||||| ||| : : ||||| ||| : : ||||| ||| : : ||||| ||| : : |||||
Db 463 RLIIIGVQSEPLFATTTAENIRYGRDVTDEMERARAEANAYDFIMEFPKFNVLVGE 522
OY 459 KGAQMSGGOKRIATARALVNRPKILLIDEATSAIDSEKSAVQAALDK 507
    : : ||||| ||| : : ||||| ||| : : ||||| ||| : : ||||| ||| : : |||||
Db 523 RGAQLSGGOKRIATARALVNRPKILLIDEATSAIDSEKSAVQAALDK 571

RESULT 5
multidrug resistance protein 1 - Chinese hamster
N:Alternate names: P-glycoprotein pgp1
C:Species: Cricetus griseus (Chinese hamster)
C>Date: 31-Dec-1990 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
R:Accession: A38696; C38696; B38696; A27126; S33768; 152823
J:Devine, S.E.; Hussain, A.; Davide, J.P.; Melera, P.W.
A:Title: Full length and alternatively spliced pgp-1 transcripts in multidrug-resistant
A:Reference number: A38696; MUID:91154265
A:Accession: A38696
A:Molecule type: mRNA
A:Residues: 1-1276 <DEV>
A:Cross-references: GB:M59253; NID:g191154; PIDN:AAA37004.1; PID:g191155
A:Accession: C38696
A:Molecule type: mRNA
A:Residues: 108-1276 <DE1>
A:Cross-references: GB:M59254; NID:g191156; PIDN:AAA37005.1; PID:g191157
A:Experimental source: clone ADX185
A:Accession: B38696
A:Molecule type: mRNA
A:Residues: 1-32,771-1276 <DE2>
A:Cross-references: GB:M59252; NID:g191152; PIDN:AAA37003.1; PID:g191153
A:Experimental source: clone ADX124
R:Endicott, J.A.; Juranka, P.F.; Sarangi, F.; Gerlach, J.R.; Deuchars, K.L.; Ling, V.
Mol. Cell. Biol. 7, 4075-4081, 1987
A:Title: Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chinese h
A:Reference number: A27126; MUID:88122132
A:Accession: A27126
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 706-1276 <END>
A:Cross-references: GB:M17897; NID:g191158; PIDN:AAA37006.1; PID:g191159
R:Zastawny, R.L.; Ling, V.
Biochim. Biophys. Acta 1173, 303-313, 1993
A:Title: Structural and functional analysis of 5' flanking and intron 1 sequences of the
A:Reference number: S33768; MUID:93305724
A:Accession: S33768
```

```
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-21 <ZAS>
A:Cross-references: EMBL:L03286
R:Teeter, L.D.; Eckersberg, T.; Tsai, Y.; Kuo, M.T.
Cell Growth Differ. 2, 429-437, 1991
A:Title: Analysis of the Chinese hamster P-glycoprotein/multidrug resistance gene pgp
A:Reference number: 152823; MUID:92088970
A:Accession: 152823
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-21 <RES>
A:Cross-references: GB:S81975; NID:g240862
A:Comment: This is an integral membrane protein overproduced in multidrug-resistant c
structurally and functionally unrelated lipophilic antitumor drugs.
C:Genetics:
A:Gene: pgp1
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: alternative splicing; ATP; duplication; glycoprotein; nucleotide binding;
F:407-601/Domain: ATP-binding cassette homology <ABC1>
F:424-431/Region: nucleotide-binding motif A (P-loop)
F:548-552/Region: nucleotide-binding motif B
F:1050-1246/Domain: ATP-binding cassette homology <ABC2>
F:1057-1074/Region: nucleotide-binding motif A (P-loop)
F:1193-1197/Region: nucleotide-binding motif B
F:87,91,96/Binding site: carbonylate (Asn) (covalent) #status predicted
F:430/Binding site: ATP (Lys) #status predicted
F:1073/Binding site: ATP (Lys) #status predicted

Query Match 60.3%; Score 1564; DB 1; Length 1276;
Best Local Similarity 56.8%; Pred. No. 1.7e-102;
Matches 301; Conservative 100; Mismatches 99; Indels 30; Gaps 4;

OY 1 MILGTLASLVNGACLPMPVLVLEMSDNLISGLVQTN----- 38
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 50 MLVGTAAIHIGVALPLMLLVFGDMTDFASVGNIPYNNATQVNASDIFGKLEEMT 109
OY 39 TYSEFRLTYVYGIVAAALIFGYIOISLMTITPAROTKRIRKOFPHSVLAODIGFSDCI 98
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 YAY-----YTGIGAGVLIVATVGSFMCCLAAGROIHKIRQFHAIMNOEIGFMDVHD 164
OY 99 IGEINTRMT-DIDKISDIDGDKIALLFONMSFFSIGLAAGLVGKMLTLVLTSTSPILMA 157
    : : ||||| ||| : : ||||| ||| : : ||||| ||| : : ||||| ||| : : |||||
Db 165 VGEINTRLTDVDSKINEGIDGKIGMFQAMATPFEGFIIIGFTRGKMLTLVILAIISPVLG 224
OY 158 SAACSRMVISLTSKLSATSKAGAAVEVLSIRTVIAFRAOEKLOXYTQMLKAKDF 217
    | : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 225 SAGIWKILISSEFDKELQAYAKAGAAVEVLAIRTVIAFGGOKKELERNNNLEBAKRL 284
OY 218 IKRTASKVSLGAVYFFMNGTGYLAFWYGTSLIINGEPGYTIGTVLAFFSVHSYCSY 277
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 GIKKATINISKAAPALLIYASVALAFWYGTSLVYSKE--YSTGVLYVFFSVLTAFSVG 342
OY 278 GAAVPHETFAIARGAAPHIPOVIDKKPSIDNESTAGYKPESTIEGTEVERKNVSFNPSR 337
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 343 QASPMIEAFANARGAAYEFIINIDNKPSIDSFSGKHKPDNIOGMLERKNHFSYPSRK 402
OY 338 SIKILKGLNLRKISGETVALVGNLNGSGKSTVQQLRLDLPDGFIMVDENDIRALNVRY 397
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 DVQIILKGLNLRKYSQGVVALVGNLNGSGKSTVQQLMQLRPLDGMVSDGQDIRTNVRY 462
OY 398 YVDHIGVVSQEPVLFGTTSNNIKRYGRDVTDEMERARAEANAYDFIMEFPKFNVLVGE 457
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 463 RLIIIGVVSQEPVLFATTTAENIRYGRDVTDEMERARAEANAYDFIMEFPKFNVLVGE 522
OY 458 EKGQMSGGOKRIATARALVNRPKILLIDEATSAIDSEKSAVQAALDK 507
    : : ||||| ||| : : ||||| ||| : : ||||| ||| : : ||||| ||| : : |||||
Db 523 ERGAQLSGGOKRIATARALVNRPKILLIDEATSAIDSEKSAVQAALDK 572

RESULT 6
555692
```





```

OY 281 VPHETFAVGAAMFHTFYOYIDKKPSIDNSTGTYGPESTIEGVKXNFVPSRPSK 340
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 351 APCIDAFANMARGAAYVFYDIIIDNNPKIDSEFGKHPDSIKGULEFNVDHSTPSRANK 410

OY 341 ILKGLNLRKSGETVALVGLNGSGKSTVVOLORLDPPDGFIMVENDIRALANRHRD 400
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 ILKGLNKGVSQGVVALVSGSGCKSTVVOQLRQVDDPGTINDQIRMFNVNLYRE 470

OY 401 HIGVNSOEPLYFTTISNNNTKYGRDVTDEEMRAAREANAIDEIMEFNNKNTILYGEK 460
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 IIGVNSOEPLYFTTISNNNTKYGRDVTDEEMRAAREANAIDEIMEFNNKNTILYGEK 530

OY 461 AOMSGGOKORAIARALVNRPKIILIDEATSLDSEKSAVOALAEK 507
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 531 AOMSGGOKORAIARALVNRPKIILIDEATSLDSEKSAVOALAEK 577

```

RESULT 8  
148123  
P:glycoprotein isoform III - Chinese hamster  
C:Species: Crictetus griseus (Chinese hamster)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 02-Feb-2001  
C:Accession: 148123  
R:Endicott, J. A.; Sarangl, F.; Ling, V.  
DNA Seq. 2, 89-101, 1991  
A:Title: Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene family  
A:Reference number: 148121; MUID:92135896  
A:Accession: 148123  
A:Status: preliminary; translated from GR/EMBL/DBD  
A:Molecule type: mRNA  
A:Residues: 1-1281 <RES>  
A:Cross-references: GR:M00042; NID:g191168; PIDN:AAA68885.1; PID:g191169  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop  
F:412-606/Domain: ATP-binding cassette homology <ABC1>  
F:429-436/Region: nucleotide-binding motif A (P-loop)  
F:1054-1250/Domain: ATP-binding cassette homology <ABC2>  
F:1071-1078/Region: nucleotide-binding motif A (P-loop)

Query Match	56.5%	Score 1464	DB 2	Length 1281
Best Local Similarity	53.5%	Pred No. 2.1e-95		
Matches 263	Conservative 108	Mismatches 108	Indels 30	Gaps 5

```

0Y 1 MILGLSLVNGACPIPLVLPGVMSNDLISGCLVQNTSF-----FRLL-----47
Db 57 MULGTIMAINANGSLPLMIVYFGEMTKEFVN-----NAGFSLPYAFSLSMINPGRILE 110
0Y 48 -----YYVGIGVALLIFGYIOISLMTITTAQOTKRIKROFHSVLAODIGFSDCDI 99
Db 111 BEMTEYAYAYYSGLSGGVLYAAVYIOVSFWTLNAGROIKKIRONFPHALLRBOEMGFDIKGT 170
0Y 100 GELNTRMT-DIDKTSIDGIGDXIALLPQMSFESGLAVGLKGGKLLTVLSTSPILMAS 158
Db 171 TELNRLRLDDLSKISSEGIGDXVGAFQAVAFEPFAGTIVGLFRGKLLTVLMAISPILGIS 220
0Y 159 AAASRMVYISLTSKELASVAKAGVAEEVLSIRITVIAFRAOEKLORYONTLKDADFC 218
Db 231 AAVMAKILSTFSDKELAYAKAGVAEBALCAITVIAFGQNNKEERYOKHLENMAKKIG 290
0Y 219 IKRTIASVSLGAVYEFPMNGSYGLAFWYGISLLINGEPCYITGLVAFFSVIHSYSCIG 278
Db 291 IKKAISANISMGIAFLLIYASAYALAFWYGSTLVTSKE--YTGNAAMTVFSLIGAFSVG 348
0Y 279 AAVPHEFEFALARGAAPHIIPVIDKKPSIDNFSTAGIKRPEISIECTVEFKNVSPYRPS 338
Db 349 QAAPIIDAFANARGAAYVIFPIDNNPKIDSFSERGHKPDSIKENLDFSDVHSYSPERAN 408
0Y 339 IKIKGLMLRKISGETVVALNGSNGSTVQVOLLORLYDPDGDGFIWMDENDIRALNRYH 398
Db 409 IKIKGLMLKVOSSGOTVALVNSGCGKTTTLQOLRLYDLPPEGITISIDGODIRENRNYL 468

```

[illegible]

RESULT 9  
DMS2  
multidrug resistance protein 2 - mouse  
N:Alternative names: P-glycoprotein MDR2  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1990 #sequence-revision 31-Dec-1990 #text-change 19-Jan-2001  
C:Accession: A30409; #70711  
R:Gros, P.; Raymond, M.; Bell, J.; Housman, D.  
Mol. Cell. Biol. 8, 2770-2778, 1988  
A:Title: Cloning and characterization of a second member of the mouse mdr gene family  
A:Reference number: A30409; MUID:86302195  
A:Accession: A30409  
A:Molecule type: mRNA  
A:Residues: 1-1276 <NS>  
A:Cross-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428  
R:Krischner, L.S.  
Nucleic Acids Res. 24, 2829-2834, 1996  
A:Title: De novo generation of simple sequence during gene amplification.  
A:Reference number: S70711; MUID:96313253  
A:Accession: S70711  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 43-92 <KTR>  
A:Cross-references: EMBL:U46839; NID:g1228142; PIDN:AAC52722.1; PID:g1228143  
A:Note: the nucleotide sequence was submitted to the EMBL Data library, January 1996  
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c  
turally and functionally unrelated lipophilic antitumor drugs.  
C:Genetics:  
A:Gene: mdr2  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane  
F:1-637, 653-1276/Region: duplication  
F:409-603/Domain: ATP-binding cassette homology <ABC1>  
F:426-433/Region: nucleotide-binding motif A (P-loop)  
F:550-554/Region: nucleotide-binding motif B  
F:1049-1245/Domain: ATP-binding cassette homology <ABC2>  
F:1067-1074/Region: nucleotide-binding motif A (P-loop)  
F:1192-1196/Region: nucleotide-binding motif B  
F:88, 94/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:433/Binding site: ATP (Lys) #status predicted  
F:1072/Binding site: ATP (Lys) #status predicted

Query Match	1463	Score	1463	DB 1	length	1276			
Best Local Similarity	53.9%	Pred.	No. 2.4e-95						
Matches	265	Conservative	103	Mismatches	111	Indels	30	Gaps	57

[illegible]





T19982  
hypothetical protein C47A10.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text-change 31-Jan-2000  
C:Accession: T19982; T24202  
R: Basham, V.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: 219207  
A:Accession: T19982  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1294 <NTL>  
A:Cross-references: EMBL:Z81484; PIDN: CAB03973.1; GSPDB: GNO0023; CESP: C47A10.1  
A:Experimental source: clone C47A10  
R: Basham, V.  
submitted to the EMBL Data Library, April 1997  
A:Reference number: 219853  
A:Accession: T24202  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1294 <NTL>  
A:Cross-references: EMBL:Z93782; PIDN: CAB07855.1; GSPDB: GNO0023; CESP: C47A10.1  
A:Experimental source: clone R12G8  
C:Genetics:  
A:Gene: CESP:C47A10.1  
A:Map position: 5  
A:Insertions: 15/2; 67/3; 123/3; 168/2; 509/1; 649/3; 826/2; 1089/3; 1158/3; 1250/1  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 44.7%; Score 1158; DB 2; Length 1294;  
Best Local Similarity 45.0%; Pred. No. 9.4e-74;  
Matches 235; Conservative 95; Mismatches 174; Indels 18; Gaps 4;

QY 1 MILGIASLVNACGLPLMLPLVLEMSDNLIS--GCLVQTN-----TYSFRL 45  
DB 46 LAVGIYVSCAGVGLPLMSIITGNVSONPVTLCITFLDPNVSASKAARAARAFESHEVQN 105  
QY 46 TLYYVIGVVAALIFGYIOISLMTTAARQTKRIRKOFHSHVLAQDIGWDSDDLNLN 105  
DB 106 CLKVYLLCGIFPAAGFIDASCFCMVCEKLSNFRROFHSVROEIAMVDKMTSGTILSNK 165  
QY 106 MTD-IDKISDGIKIALIFQMSFSGIAGLVKGMKLTLYTSTPLIMASAACSR 164  
DB 166 LPDNLERYREGDVGIAFQMAQF IGFPAAFTYDMLTILMISLSPFMMICGLFLAK 225  
QY 165 WISLTSKELSAVSKAGAAVEVLSSIRTVIAFRAQEKELORYTONLKDADFGIKRTIA 224  
DB 226 LLATATATREAKOYAAVAGIAIEVLISIRTVIAFNGOEYCKRYEDALEHKGRTGKIKSL 285  
QY 225 SKVSLGAVYFFPMNGTYGLAFWYGTSLINGEFGYITGVLAFFSVIHSSYICGAAPPH 284  
DB 286 IGAGIASFSEVLIYASYCLAFWGTNFEVSGR--LESGLVLTVEFVSVMGSMALGQAGQOF 343  
QY 285 EEFPAIARAAAFHFOVIDKPSIDNFSAGYKPSIEGTVKKNVSFNPSPSKIKILKG 344  
DB 344 ATIGTALGAASLVIEYDIPIDAYSTEGOTPSKISGRISVKNKYEFYPTADYKILKG 403  
QY 345 LNLRIKSETVALVGLNGSGKSTVQVQLRLYDPDGFIMVDENDIRALNVHRHYDHICV 404  
DB 404 VSLDAQPGQYVALVGVSSGCKSTIIQLQRFYNPDAGQILDDIPIDENIKIYLLQVGV 463  
QY 405 VSGEVLFGTTISNNIKYGRDDYTDDEMERARAENAYDITMEFPNKENTLVGEKGAQMS 464  
DB 464 VSGEPLNFTSIEQNIIRGRSDVSDEDIARALKENANADFIKTFEGGLTIVLGDGVQMS 523  
QY 465 GSGOKORIAIARALVNPKILLDEATSLDSEKSAVOALE 506  
DB 524 GSGOKORIAIARALVNPKILLDEATSLDSEKSAVOALE 565

RESULT 15  
T31073

multidrug resistance P-glycoprotein - nematode (Haemonchus contortus)  
C:Species: Haemonchus contortus  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text-change 11-May-2000  
C:Accession: T31073  
R: Xu, M.; Molento, M.; Blackhall, W.; Ribeiro, P.; Beech, R.; Prichard, R.  
Mol. Biochem. Parasitol. 91, 327-335, 1998  
A:Title: Ivermectin resistance in nematodes may be caused by alteration of P-glycopro  
A:Reference number: 220968; MUID:98226176  
A:Accession: T31073  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1275 <XUM>  
A:Cross-references: EMBL:AF003908; NID: g3057041; PID: g3057042; PIDN: AAC38987.1  
A:Experimental source: developmental stage adult  
C:Genetics:  
A:Gene: PGP-A  
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

Query Match 40.5%; Score 1050.5; DB 2; Length 1275;  
Best Local Similarity 42.6%; Pred. No. 3.7e-66;  
Matches 223; Conservative 94; Mismatches 186; Indels 21; Gaps 5;

QY 4 GILASLVNACGLPLMLPLVLEMSDNLIS--GCLVQTN-----TYSFRL 44  
DB 54 GILACVHAGSAGSVIGIYVGLGMTTFELRNQNSEFVLGTVSRDEGLPALTKREFDTLVR 113  
QY 45 LFLYVIGVVAALIFGYIOISLMTTAARQTKRIRKOFHSHVLAQDIGWDSDDLNLN 104  
DB 114 YCLYVIGLGFAMFANSYIOVCEMFAERITHKRLKRYLAKILRQOISWFIQNGNLN 173  
QY 105 RMT-IDKISDGIKIALIFQMSFSGIAGLVKGMKLTLYTSTPLIMASAACSR 163  
DB 174 RLTDLERYREGDVGIAFQMAQF IGFPAAFTYDMLTILMISLSPFMMICGLFLAK 225  
QY 164 RMT-IDKISDGIKIALIFQMSFSGIAGLVKGMKLTLYTSTPLIMASAACSR 164  
DB 224 KIVATRTVOEQTVAAGIAIEVLISIRTVIAFNGOEYCKRYEDALEHKGRTGKIKSL 285  
QY 224 ASKVSIGAVYFFPMNGTYGLAFWYGTSLINGEFGYITGVLAFFSVIHSSYICGAAPPH 284  
DB 294 YMGVGVGFCQMCYVSYALAFWYGSVLIN--DPALDRGRIFFPFAVMGSAALGCLPH 352  
QY 284 EEFPAIARAAAFHFOVIDKPSIDNFSAGYKPSIEGTVKKNVSFNPSPSKIKILKG 343  
DB 353 LNTISIAAGVAVSVIINSRKPIDYSLDGIVLNMMRGSIRKKNVHFSYPSRRLOILK 412  
QY 344 GLNLRKSETVALVGLNGSGKSTVQVQLRLYDPDGFIMVDENDIRALNVHRHYDHICV 403  
DB 413 GVSLOVSGQKIALVGVSSGCKSTVNNLLRFDYDPRGKVTITDDIDVCDLNOKLREQIG 472  
QY 404 VSGEVLFGTTISNNIKYGRDDYTDDEMERARAENAYDITMEFPNKENTLVGEKGAQMS 463  
DB 473 VSGEPLNFTSIEQNIIRGRSDVSDEDIARALKENANADFIKTFEGGLTIVLGDGVQMS 523  
QY 464 GSGOKORIAIARALVNPKILLDEATSLDSEKSAVOALE 506  
DB 533 GSGOKORIAIARALVNPKILLDEATSLDSEKSAVOALE 565

Search completed: April 22, 2002, 14:35:58  
Job time: 190 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:37:44 ; Search time 24.77 Seconds

(without alignments)  
760.828 Million cell updates/sec

Title: US-09-873-409-8

Sequence: 1 MILGTLASLVGACLPMLPL.....SEKSAVQALEKOPRYSF 514

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : swissprot\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1590	61.3	1276	MDR1_MOUSE	P06795 mus musculi
2	1579	60.9	1280	MDR1_HUMAN	P08183 homo sapien
3	1578	60.9	1276	MDR2_CRIGR	P21449 cricetulus
4	1571.5	60.6	1277	MDR1_RAT	P43245 rattus norv
5	1565.5	60.4	1276	MDR3_MOUSE	P21447 mus musculi
6	1565	60.4	1276	MDR1_CRIGR	P21448 cricetulus
7	1479	57.1	1279	MDR3_HUMAN	P21439 homo sapien
8	1464	56.5	1281	MDR3_CRIGR	P23174 cricetulus
9	1463	56.4	1276	MDR2_MOUSE	P21440 mus musculi
10	1445	55.7	1278	MDR2_RAT	P08201 rattus norv
11	1273.5	49.1	1321	AB11_HUMAN	O95342 homo sapien
12	1243.5	48.0	1321	AB11_RAT	O70127 rattus norv
13	1227.5	47.4	1321	AB11_MOUSE	O99430 mus musculi
14	1176.5	45.4	1321	AB11_RABIT	O99073 oryctolagus
15	1025.5	39.6	1302	MDR4_DROME	O00449 drosophila
16	1021	39.4	1321	MDR1_MOUSE	P34712 mus musculi
17	972.5	37.5	1324	MDR3_MOUSE	P34713 caenorhabdi
18	945.5	36.5	1362	MDR1_MOUSE	P36619 schizosacch
19	935	36.1	1280	MDR1_MOUSE	O06034 leishmania
20	923	35.6	1302	MDR5_DROME	O00748 drosophila
21	699	27.0	735	ABCB_HUMAN	O99072 homo sapien
22	696.5	26.9	820	MDL2_YEAST	P33311 saccharomyc
23	675	26.0	1419	MDR1_MOUSE	P13568 plasmodium
24	669.5	25.8	695	MDL1_YEAST	P33310 saccharomyc
25	663.5	25.6	685	MDL1_MOUSE	P97998 candida alb
26	656.5	25.3	703	TAP2_RAT	P36372 rattus norv
27	645.5	24.9	726	YFYE_SCHPO	O99747 schizosacch
28	641	24.7	686	TAP2_HUMAN	O03519 homo sapien
29	635.5	24.5	702	TAP2_MOUSE	P36371 mus musculi
30	622	24.0	724	TAP1_MOUSE	P21958 mus musculi
31	619.5	23.9	725	TAP1_RAT	P36370 rattus norv
32	602	23.2	582	MSBA_ECOLI	P27299 escherichia
33	598	23.1	748	TAP1_HUMAN	O03518 homo sapien

## ALIGNMENTS

RESULT	ID	MDR1_MOUSE	STANDARD	PRT	1276 AA.
AC	P06795	MDR1_MOUSE			
DT	01-JAN-1988 (Rel. 06, Created)				
DT	01-JAN-1988 (Rel. 06, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).				
GN	ABCB1 OR PGY1 OR PGY1-1 OR MDR1 OR MDR1B.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87028229; PubMed=3768958;				
RA	Gros P., Croop J., Housman D.;				
RT	"Mammalian multidrug resistance gene: complete cDNA sequence				
RT	indicates strong homology to bacterial transport proteins.";				
RL	Cell 47:371-380(1986).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89367274; PubMed=2570420;				
RA	Raymond M., Gros P.;				
RT	"Mammalian multidrug resistance gene: correlation of exon				
RT	organization with structural domains and duplication of an ancestral				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).				
RN	[3]				
RP	SEQUENCE OF 1-21 FROM N.A.				
RX	MEDLINE=91042535; PubMed=2248681;				
RA	Raymond M., Gros P.;				
RT	"Cell-specific activity of cis-acting regulatory elements in the				
RT	Promoter of the mouse multidrug resistance gene mdr1.";				
RL	Mol. Cell. Biol. 10:6036-6040(1990).				
CC	- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED				
CC	DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.				
CC	- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC	- PPM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE				
CC	LINKER DOMAIN.				
CC	- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE				
CC	RELATED BUT DISTINCT CELLULAR GENES.				
CC	- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY				
CC	(ABC TRANSPORTERS). MDR SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: M14757; AAA79005.1; -				
DR	EMBL: M60348; AAA39513.1; -				
DR	PIR: A33719; DVM51.				

O9WC4 thermotoga  
P44407 haemophilus  
P54719 bacillus su  
P2638 anabaena sp  
Q2868 dictyostell  
P5469 rhizobium s  
P12866 saccharomyc  
P47261 mycoplasma  
P10089 escherichia  
P08716 escherichia  
P11599 proteus vul  
P26760 actinobacil



or send an email to license@slb-slb.ch).

CC -----  
 DR EMBL: M14758: AAAS9575.1; -  
 DR EMBL: M29447: AAAS9576.1; -  
 DR EMBL: M29424: AAAS9576.1; JOINED.  
 DR EMBL: M29425: AAAS9576.1; JOINED.  
 DR EMBL: M29426: AAAS9576.1; JOINED.  
 DR EMBL: M29427: AAAS9576.1; JOINED.  
 DR EMBL: M29428: AAAS9576.1; JOINED.  
 DR EMBL: M29429: AAAS9576.1; JOINED.  
 DR EMBL: M29430: AAAS9576.1; JOINED.  
 DR EMBL: M29431: AAAS9576.1; JOINED.  
 DR EMBL: M29432: AAAS9576.1; JOINED.  
 DR EMBL: M29433: AAAS9576.1; JOINED.  
 DR EMBL: M29434: AAAS9576.1; JOINED.  
 DR EMBL: M29435: AAAS9576.1; JOINED.  
 DR EMBL: M29436: AAAS9576.1; JOINED.  
 DR EMBL: M29437: AAAS9576.1; JOINED.  
 DR EMBL: M29438: AAAS9576.1; JOINED.  
 DR EMBL: M29439: AAAS9576.1; JOINED.  
 DR EMBL: M29440: AAAS9576.1; JOINED.  
 DR EMBL: M29441: AAAS9576.1; JOINED.  
 DR EMBL: M29442: AAAS9576.1; JOINED.  
 DR EMBL: M29443: AAAS9576.1; JOINED.  
 DR EMBL: M29444: AAAS9576.1; JOINED.  
 DR EMBL: M29445: AAAS9576.1; JOINED.  
 DR EMBL: M29446: AAAS9576.1; JOINED.  
 DR EMBL: AF016355: AAB69423.1; -  
 DR EMBL: AC002457: AAC82531.1; -  
 DR EMBL: M37724: AAA8047.1; -  
 DR EMBL: M37725: AAA8048.1; -  
 DR EMBL: X58723: CAA41558.1; -  
 DR PIR: A25059: DVH01.  
 DR PIR: A34914: A34914.  
 DR MIM: 171050: -  
 DR InterPro: IPR003593: AAA.  
 DR InterPro: IPR003439: ABC\_transporter.  
 DR InterPro: IPR001140: ABC\_transporter\_tmemb.  
 DR InterPro: IPR001687: ATP\_GTP\_A.  
 DR Pfam: PF00664: ABC\_membrane; 2.  
 DR Pfam: PF00005: ABC\_tran; 2.  
 DR SMART: SM00382: AAA; 2.  
 DR PROSITE: PS00211: ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
 KM Multigene family.  
 FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 52 72 POTENTIAL.  
 FT TRANSMEM 120 140 POTENTIAL.  
 FT TRANSMEM 189 209 POTENTIAL.  
 FT TRANSMEM 216 236 POTENTIAL.  
 FT TRANSMEM 297 317 POTENTIAL.  
 FT TRANSMEM 326 346 POTENTIAL.  
 FT DOMAIN 347 710 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 711 731 POTENTIAL.  
 FT TRANSMEM 757 777 POTENTIAL.  
 FT TRANSMEM 833 853 POTENTIAL.  
 FT TRANSMEM 854 874 POTENTIAL.  
 FT TRANSMEM 937 957 POTENTIAL.  
 FT TRANSMEM 974 994 POTENTIAL.  
 FT DOMAIN 995 1280 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 427 434 ATP (BY SIMILARITY).  
 FT NP\_BIND 1070 1077 ATP (BY SIMILARITY).  
 FT REPEAT 1 637  
 FT REPEAT 638 1280  
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 23 23 S -> V (IN REF. 6).  
 FT CONFLICT 185 185 G -> V (IN REF. 1 AND 3).  
 FT CONFLICT 336 336 MISSING (IN REF. 3).  
 FT CONFLICT 412 412 G -> A (IN REF. 3).  
 FT CONFLICT 438 438 O -> S (IN REF. 3).  
 SQ SEQUENCE 1280 AA: 141462 MW: 141462 MW: ABLC279531P43675 CRC64;

Query Match 60.9%; Score 1579; DB 1; Length 1280;  
 Best Local Similarity 58.1%; Pred. No. 2e-99; Indels 22; Gaps 5;  
 Matches 306; Conservative 100; Mismatches 99; Indels 22; Gaps 5;  
 1 MILGILASLVNGACLPMPVLVGENSD-----NLISGLVQV--NTYSFF----- 43  
 51 MYVGLTAIITHAGLPMLLVGEMTDIFANNGNLEDLSNTNSDINDIDGFPNMLEED 110  
 44 --RLTVYGVIGVAAALIEGYIOISLMIITTAHQTRKIRKQPHSVLAQDIGMPSDCIGE 101  
 111 MRRVYVYSGIAGVLAAYIYGVSWCLAGQIKIRKQPHFHAIMRQIFGDFVDVGE 170  
 102 LNTFRFT-DIDKTSIDGIDGDKIALLPQNMSPFSIGLAVGLKGLKTLVTLSTPLMASNA 160  
 171 LNTRLTLDVSKINEGIDGIGFQSMATFEGFVGFGRKRLTLVLAISPVLGLSNA 230  
 161 ACSRWYISLTSKELSAVSAGAAVEVLSSIRTVIAFRAQEKELQRYTONLRADQFGIK 220  
 231 VAKILISFTDKELLALYAKAGAAVEVLAIKRTVIAFGQKKLEKRYKNLEAKRIGIK 290  
 221 RTIAKVSLSGAVYFPMNGTYGLAFWYGSLLINGEPGYITGVLAFFSVIHSSYCIGA 280  
 291 KAITANISIGAFLLIYASVALAFWYGTVLVSGE--YSIGVLFVFFSVLIGAFSVGQA 348  
 281 VHFETPAIARCAARHIFQVYIDKPSIDNFSIAGYKPSIECTVEKRVSNFSPKPSIK 340  
 349 SPSEIAFANARAAVEIFIIDNKPSIDSYSGHKNPDKNLEKRNHFSPKREYK 408  
 341 ILKGLNLRKSGEYVALVGLNGSKSTVQVQLQRLYPDPDGFIIVDENDIRALNVNRYD 400  
 409 ILKGLNLKVSQGVYALVNGSCGKSTVQVQLQRLYPDPEGAVSDGQIRIRINVFLRE 468  
 401 HIGVVSQEPVLEFTGISNNIKYGRDVTDEMERARANAAYDFIMEPNKNTLVGEGK 460  
 469 IIGVVSQEPVLEFTGISNNIKYGRDVTDEMERARANAAYDFIMEPNKNTLVGEGK 528  
 461 AOMSGGOKORIAARALVNPRLILIDETSAIDSKSAVQAALKE 507  
 529 AOLSGGOKORIAARALVNPRLILIDETSAIDSKSAVQAALKE 575  
 RESULT 3  
 MDR2\_CRIGR ID MDR2\_CRIGR STANDARD: PRT; 1276 AA.  
 AC P21449;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).  
 GN PGY2 OR PGY2.  
 OS Cricetulus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus;  
 OX NCBI\_Taxid=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92135896; PubMed=1685679;  
 RA Endicott J.A., Sarangi F., Ling V.;  
 RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein  
 RT gene family.";  
 RL DNA Seq. 2:89-101(1991).  
 [2]  
 RP SEQUENCE OF 622-1276 FROM N.A.  
 RX MEDLINE=88122132; PubMed=2893255;  
 RA Endicott J.A., Juranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,  
 RA Ling V.;  
 RT "Simultaneous expression of two P-glycoprotein genes in  
 RT drug-sensitive Chinese hamster ovary cells.";  
 RL Mol. Cell. Biol. 7:4075-4081(1987).  
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED



```

CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60041; AAA68884.1; -.
DR EMBL; M17896; AAA37007.1; -.
DR PIR; B27126; DVHY2C.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001140; ABC_transportr_tmem.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 708 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 831 851 POTENTIAL.
FT TRANSMEM 852 872 POTENTIAL.
FT TRANSMEM 935 955 POTENTIAL.
FT TRANSMEM 972 992 POTENTIAL.
FT DOMAIN 993 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (POTENTIAL).
FT NP_BIND 1068 1075 ATP (POTENTIAL).
FT REPEAT 1 635
FT REPEAT 636 1276
SQ SEQUENCE 1276 AA; 141057 MW; 5096B1385628812D CRC64;
Query Match 60.98; Score 1578.5; DB 1; Length 1276;
Best Local Similarity 57.48; Pred. No. 2,1e-99;
Matches 305; Conservative 99; Mismatches 96; Indels 31; Gaps 5;
OY 1 MILGLIATLVNACGLPLPLVIGMSDN-----LISGLTVQ-T 37
DB 51 MVLGTLAVLVHGTSLPLMLVFGNMTDSFKAERSIMPNMTNOSINTTEVYISGLEEDM 110
OY 38 NTYSFFRLTLVYVIGVALAFGYQIQLSWITTAARQTKRINKQFHSVLAQDIDGPDSC 97
DB 111 ATYAV-----YTGIGAGCVLIVAYIQVSEFWCLAAGRQINKRKQFHAIMNDEIGMFDVH 165
OY 98 DIGELNMT-DIDKISGIDGKTALLFQNNSTFISGLAVGLVKGWKLTVLTLSPLIM 156
DB 166 DIGELNMTLTDVSKINGIDBKIGMFQSIATFLAIFVIGISGMKLTLLVILAVSPILG 225
OY 157 ASAACSRWVSLTSKESLAYSAGAAVEEVLSSRTVIAFAQKEORTONKDKD 216
DB 226 LSSAMMAVVLVSFTKELQATKAGAAVEEVLATRTVIAAGGKKELERYKNLEAKN 285
OY 217 FGKRTIASKVSLSGAVYEFPMNGTYGLAFWYGTSLILNGEPGYTIGTVLAVFEVSIISSVC 276

```

```

DB 286 VGKRAVTANISIGIAVLVYASALAFWYGTSLVLSNE--YSVGVLTFEFSILFGTFS 343
OY 277 IGAVPHPEPAIRGAFAHIFQVIDKKPSIDNSTAGYKPESTGEIVPEKFNFSNPSR 336
DB 344 IGHIAPILEVPAANRGAAYELFKIIDNPSIDSTQGHKPDSTMGNLEFRNHFSPSR 403
OY 337 PSKILGLNLRKSGETVALVGLNGSKSTVYOLLRLXDPDDGFTMVDENDIRALNVR 396
DB 404 SGKILGLNLKVGSGVVALVKGSGKSTVYOLLRLXDPPTGEGVVISDQDIRITNVR 463
OY 397 HYRPHIGVSEPEVLFGTTSNNIKYGRDVTDEMEAREANAPFIMEFPKFTLV 456
DB 464 YLRLEIGVSEPEVLFATTAENIRYGRNVTMDIEKAAVEANAAYDFIMKLPKFDTLV 523
OY 457 GEGKMGSGGOKRATARALVRNPKIILDEPATSALDSEKSAVOALEK 507
DB 524 GERGHQLSGGOKRIATARALVRNPKIILDEPATSALDSEAVQNALDK 574

RESULT 4
MDRL_RAT STANDARD; PRT; 1277 AA.
ID MDRL_RAT
AC P43245;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).
GN ABCB1 OR PGL1 OR MDRL OR MDR1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92039081; PubMed-1682220;
RA Silverman J.A., Raunio H., Gant T.W., Thorngelsson S.S.;
RT "Cloning and characterization of a member of the rat multidrug
RT resistance (mdr) gene family.";
RL Gene 106:229-236(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
CC LINKER DOMAIN.
CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
CC RELATED BUT DISTINCT CELLULAR GENES.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81855; -. NOT_ANNOTATED_CDS.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001140; ABC_transportr_tmem.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 71 POTENTIAL.
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.

```



```

FT NE_BIND 1066 1073 ATP (BY SIMILARITY).
FT REPEAT 1 632
FT CARBOHYD 633 1276
FT CARBOHYD 83 83
FT CARBOHYD 87 87
FT CARBOHYD 90 90
FT CONFLICT 526 527
FT CONFLICT 939 939
FT CONFLICT 1036 1036
FT SEQUENCE 1276 AA; 140754 MW; 75C71F33E1F56481 CRC64;

Query Match 60.4%; Score 1565.5; DB 1; Length 1276;
Best Local Similarity 56.5%; Pred. No. 1.6e-98;
Matches 299; Conservative 103; Mismatches 98; Indels 29; Gaps 4;

OY 1 MILDILASLVNGACLPMLPLVLEKSDNLISGLVQTN-----T 39
DB 50 MLVGTALAIHGVALPLMLLIFGDMTDSFASGVNKSNTNMSADKRAMFAKLEEMTT 109
OY 40 YSFRLTLVYVIGVAAALIFGIDISLIIITAAROTAKRIKROFHSVLAODIGMFDSDI 99
DB 110 YAY-----YYTGAGVLTIVAYIOVSEWCIAAGROHRIKROKFEHALMNOEIGMFDVADV 164
OY 100 GELNTRMT-DIDKISDGDIGDKIALLFQMSFYSIGLAVGLKWKGLTLVLTSTPLIMAS 158
DB 165 GELNTRLTDVSKINIEGDKIGMFFQMAITFFGFIIGFRKMLTLVLAISPVGLS 224
OY 159 AAACSRMVISLTSKELASRAGAAVEVLSIRTVIAFRAOEKELORYTONLKADKDFG 218
DB 225 AGIWAIIKSTFDEKLHAKAGAAVEVLAIRTVIAFGQKELERYNNLNLEAKRLG 284
OY 219 IKRTIASVSGAVYFENGTYGLAFWGTGSLIINGEPGRTIGTVLAFESVHSSVYIG 278
DB 285 IKKATYANISMGAAFLTLIYASALAFWGTSLVLSKE--YSIGVLTVEFVLGAESVG 342
OY 279 AAVPHFEFALARGAFAHIFQVIDKPSIDNFSTAGYKPEISIEGVFEKNSFNPSPRS 338
DB 343 QASPIIEFANARGAAYEVKIIIDNKPSIDSPSKSGHNPDIQGLERKNHFSPSKE 402
OY 339 IKILGLMLRKSGETVALVGLNCGSKSTVQVQLRLYDPDGFIMVDENIRALNRYH 398
DB 403 VOILKGLMLKYSGQTVLVNCGSKSTVQVQLRLYDPDGMVSDIGODIRITNVRYL 462
OY 399 RDHGVVSGQEVLTGRTTSSNIIKRGDQVDEEMERARERANAVYFIEMFNKFTLVGE 458
DB 463 REIIGVVSQEVLPATYTAENIRYGRDEVYDEIEKAVKANAADYFIKLPHQFDTLVGE 522
OY 459 KGAQMSGGOKORIAARALVNPVKILIDEATSSALDSKSAVQAALK 507
DB 523 RGAQMSGGOKORIAARALVNPVKILIDEATSSALDSKSAVQAALDK 571

RESULT 6
MDRL_CRIGR STANDARD; PRT; 1276 AA.
AC P21448;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 1 (P-GLYCOPROTEIN 1).
GN ABCB1 OR PGY1 OR PGP1.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OC NCBI_TaxID=10029;
OX 111
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
FT gene family.";
```

```

RL DNA Seq. 2:89-101(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91154265; PubMed=1671863;
RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;
RT "Full length and alternatively spliced pgp1 transcripts in multidrug-
RT resistant Chinese hamster lung cells."
J. Biol. Chem. 266:4545-4555(1991).
RN [3]
RP SEQUENCE OF 706-1276 FROM N.A.
RX MEDLINE=88122132; PubMed=2893255;
RA Endicott J.A., Duranka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
RA Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
RT drug-sensitive Chinese hamster ovary cells."
Mol. Cell. Biol. 7:4075-4081(1987).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- MISCELLANEOUS: PGP1 ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60040; AAA68883.1; -
DR EMBL: M59253; AAA37004.1; -
DR EMBL: M17897; AAA37006.1; -
DR PIR: A38696; DVHYIC.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding: Glycoprotein; Transport; Repeat;
KW Multigene family.
KW DOMAIN
FT FT 1 50 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 51 71
FT TRANSMEM 117 137
FT TRANSMEM 186 206
FT TRANSMEM 213 233
FT TRANSMEM 294 314
FT TRANSMEM 323 343
FT TRANSMEM 344 707
FT TRANSMEM 754 774
FT TRANSMEM 830 850
FT TRANSMEM 851 871
FT TRANSMEM 934 954
FT TRANSMEM 971 991
FT TRANSMEM 992 1276
FT DOMAIN 992 1276
FT NP_BIND 424 431
FT NP_BIND 1067 1074
FT REPEAT 1 635
FT REPEAT 636 1276
FT CONFLICT 338 339
FT SEQUENCE 1276 AA; 140925 MW; 4AF3F92A186BADFF CRC64;

Query Match 60.4%; Score 1565; DB 1; Length 1276;
Best Local Similarity 56.8%; Pred. No. 1.7e-98;
Matches 301; Conservative 100; Mismatches 99; Indels 30; Gaps 4;
```

```

OY 1 MILGILASLVNACGLPLMPVIGEMSDNLISGLVOTN----- 38
DB 50 MLVGTLLAIHGVALLPLMLVFGDMTDSFASVGNLPTNATNNAOVNASDIFGKLEEMT 109
OY 39 TYSEFRLTYVYVIGVALLFGYIDISLITTAAROTRRIRKOPFHSVLAODIGWDCSD 98
DB 110 TYAV-----YVTGIGAGVLIVAYIGVSWCLAGRQHKIRKQFPHALMNGDIGWFDVD 164
OY 99 IGEINTNMT-DIDKISDGIKIALIFONMSTFISGLAVGLVKGKMLVLTLSPLMA 157
DB 165 VGEINTRLTLDVSKINEIGDKIGMFQAMATFEFGFTIGFRGKMLTLVILAIISPVLG 224
OY 158 SAAACSRVISTSKELTSSVAKAGAVAEVISTRTVLAFAROEELORYONLMDADF 217
DB 225 SAGIWKALISSTFKELQATYAKAGAVAEVLAIPTVIAFGQKKELEKYNLLEAKRL 284
OY 218 GIKRTIASKVSLGAVYFPMNGTYGLAFWYGTSLINGEPYTGTVLAFVFSVHSYCI 277
DB 285 GIKKATINISMGAFLLIVASYALAFWYGTSLVSKS--YSIGVLYVFFAVLIGAFSI 342
OY 278 GAVPHETPALARGAAPHIQVINDKPSIDNESTAGYKPSIEGTEKFNVSFVPSRP 337
DB 343 GQASENIEAFANARGAAVEIFNIIIDNKPSIDSFSGKGYKPKDINKLEFKNIHFSYPSRK 402
OY 338 SIKITKGLNLRKSGEYALVGLNGSGSTVOLLQRLYDPDDGTIMVDENDIRALNVRH 397
DB 403 DVQILKGLNLVQSGVQVYALVNGSGSGSTVOLLQRLYDPDEGVASIDGODIRTNVRY 462
OY 398 YRDHIGVSOEPVLEGTITISNNIKYGRDDVDEEMERARAEANADFMENPKNTNVLG 457
DB 463 LREITGVVSOEPVLEPATITIAENIRKRGRENTVMDTEKAVKENADFMKLPKRDITLVG 522
OY 458 EKGAGMSGGOKORIAIARALVNPRLILDEATSALDESKSAYOALEK 507
DB 523 ERGAQLSGGOKORIAIARALVNPRLILDEATSALDESEAVOALDK 572

RESULT 7
MDR3_HUMAN STANDARD; PRT; 1279 AA.
ID MDR3_HUMAN
AC P21439;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3).
GN ABCB4 OR PGY3 OR MDR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; PubMed=2906314;
RA van der Bliek A.M., Koelman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL Gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=8811519; PubMed=2892668;
RA van der Bliek A.M., Baas F., ten Houte de Lange T., Koelman P.M.,
RA van der Velde-Koerts T., Borst P.;
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL EMBO J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; PubMed=2002063;
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
-1- FUNCTION: ENRGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED

```

```

CC CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT
CC CC CAPABLE OF CONFERRING DRUG RESISTANCE. MEDIATES THE TRANSLLOCATION
CC CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
CC CC HEPATOCYTE.
CC CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC CC -1- DISEASE: DEFECTS IN PGY3 ARE A CAUSE OF PROGRESSIVE FAMILIAL
CC CC INTRAHEPATIC CHOLESTASIS TYPE IIT (PEIC), A FORM OF AUTOSOMAL
CC CC RECESSIVE LIVER DISORDERS. CHARACTERIZED BY EARLY ONSET OF
CC CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
CC CC ADULTHOOD.
CC CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@sib-sib.ch).
CC CC -----
DR EMBL; M23234; AAA36207.1; -.
DR EMBL; X06181; CAA29547.1; -.
DR PIR; JS0051; DVH03.
DR PIR; A42213; A42213.
DR HSSP; P13569; 1NBD.
DR MIM; 171060; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001140; ABC_transporter_tmem.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 354 POTENTIAL.
FT DOMAIN 355 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.
FT TRANSMEM 854 873 POTENTIAL.
FT TRANSMEM 937 956 POTENTIAL.
FT TRANSMEM 976 993 POTENTIAL.
FT DOMAIN 994 1279 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (BY SIMILARITY).
FT NP_BIND 1069 1076 ATP (BY SIMILARITY).
FT REPEAT 1 640
FT REPEAT 641 1279
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1093 1093 V -> VVYDPGQ (IN REF. 2).
SQ SEQUENCE 1279 AA; 140682 MW; 3D56C98B5C8D6087 CRC64;
Query Match 57.1%; Score 1479; DB 1; Length 1279;
Best Local Similarity 54.1%; Pred. No. 1.2e-92;
Matches 285; Conservative 106; Mismatches 110; Indels 26; Gaps 5;
OY 1 MILGILASLVNACGLPLMPVIGEMSDNLISGLVOTNYSR---FRLTL----- 47
DB 57 MSLGITIMAHSGPLMLVIFGEMTDKFDV---TAGNFSPPVFSLSLAMPGRILEEE 112
OY 48 -----YVTGIGAGVLIVAYIGVSWCLAGRQHKIRKQFPHVLAODIGWDCSDIGE 101

```

Db 113 MFRVAYVYSGLAGVLAAYIQVSEFWTLAAGROIRKIROKFFHAILRQELIGWMDINDTTE 172  
 QY 102 LNRMT-DIDKISDGIQKALLFQNMSPFSIGLAVGVKMLTVLTSTPLIMASAA 160  
 Db 173 LNRLLDDSKISGIGDKVGFQAVTFPAGFVGTGKMLTVLIMASPLIGLSAA 232  
 QY 161 ACSRWYISLTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELORYTONMKADDFGK 220  
 Db 233 VMKILISAFSDKELAAVAKAGAAVEALGAKIRVIYAFGGONKELEROKLENAKELGK 292  
 QY 221 RTTASVSLGAVFPMNGTYGLAFWGTSLINGEPGTYTGLVLAFFSVIHSYCIGAA 280  
 Db 293 KASIANISMGIAPLLIYASALAFWYGTSLVTSKE--YTGAMATVFESILIGAFVGOA 350  
 QY 281 VPHFEPFARGAAPHFEOYIDKKPSIDNSTAGYKPESEGTVEFENVPNSPSRSIK 340  
 Db 351 APCIDAFANRGAAYVFDIIDNNPKIDSPSERGHKPDSTIKNEFDVHFSTPSRANVK 410  
 QY 341 ILKGLNLRIKSGEVALVGLNGSGKSTVVOQLQRLYDPDGFIMVDENDIRALNVHRHYD 400  
 Db 411 ILKGLNLKVGSGQVALVSGSGCKSTVQQLQRLYDPDGTINIDGQDIRNFVNYLRE 470  
 QY 401 HICVAVSOEPLVFGTTSNNIKYGRDVTDEEMERARANAADTFMEFPNKFTLVGEGK 460  
 Db 471 IICVAVSOEPLVFGTTSNNIKYGRDVTDEEMERARANAADTFMEFPNKFTLVGEGK 530  
 QY 461 AOMSGGOKORAIARALVRNPKIILDEATSAIDSEKSAVOALAEK 507  
 Db 531 AQLSGGOKORAIARALVRNPKIILDEATSAIDSEKSAVOALAEK 577

## RESULT 8

MDR3\_CRIGR STANDARD; PRT; 1281 AA.

ID MDR3\_CRIGR STANDARD; PRT; 1281 AA.  
 AC P23174;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DE 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN 3).  
 GN PGY3 OR PGY3.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92135896; PubMed=1685679;  
 RX Endicott J.A., Sarangl F., Ling V.;  
 RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein  
 RT gene family.";  
 RL DNA Seq. 2:89-101(1991).  
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT  
 CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY  
 CC CANNOT.  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M60042; AAA68885.1; -.  
 DR HSSP: P13569; INBD.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transportr.

DR InterPro: IPR00140; ABC\_transportr\_tmam.  
 DR InterPro: IPR001687; ATP\_gtp\_A.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran\_2.  
 DR SMART: SM00382; AAA\_2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 DR KAMulti-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
 KW Multigene family.  
 FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 POTENTIAL.  
 FT TRANSMEM 122 142 POTENTIAL.  
 FT TRANSMEM 191 211 POTENTIAL.  
 FT TRANSMEM 218 238 POTENTIAL.  
 FT TRANSMEM 299 319 POTENTIAL.  
 FT TRANSMEM 328 348 POTENTIAL.  
 FT DOMAIN 349 712 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 713 733 POTENTIAL.  
 FT TRANSMEM 758 778 POTENTIAL.  
 FT TRANSMEM 834 854 POTENTIAL.  
 FT TRANSMEM 855 875 POTENTIAL.  
 FT TRANSMEM 938 958 POTENTIAL.  
 FT TRANSMEM 975 995 POTENTIAL.  
 FT DOMAIN 996 1281 CYTOPLASMIC (POTENTIAL).  
 FT NF\_BIND 429 436 ATP (POTENTIAL).  
 FT NF\_BIND 1071 1078 ATP (POTENTIAL).  
 SQ SEQUENCE 1281 AA; 140866 MW; 2203EF61EBB29602 CRC64;

Query Match 56.5%; Score 1464; DB 1; Length 1281;  
 Best Local Similarity 53.5%; Pred. No. 1,3e-91;  
 Matches 283; Conservative 108; Mismatches 108; Indels 30; Gaps 5;

QY 1 MILGTLASVAGACLPLMPVLVGLGMSDNLISGLCVQNTYSF-----FRITL----- 47  
 Db 57 MLGTLMAIAHAGSLPLMIVFGEMTDKPVN-----NAGFSLPVNFSLMIMPGHILE 110  
 QY 48 -----YVVGIGVAALIFGYIQIISLMTITANQTRIRKQFHSVLAODIGFDSQDI 99  
 Db 111 EEMTRAVYVYSGLAGVLAAYIQVSEFWTLAAGROIRKIROKFFHAILRQELIGWMDINDTTE 170  
 QY 100 GELTWTMT-DIDKISDGIQKALLFQNMSPFSIGLAVGVKMLTVLTSTPLIMAS 158  
 Db 171 TELMTRTLDDISKISGIGDKVGFQAVTFPAGFVGTGKMLTVLIMASPLIGLS 230  
 QY 159 AACSRWYISLTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELORYTONMKADDFG 218  
 Db 231 AAVNAKILISTSDKELAAVAKAGAAVEALGAKIRVIYAFGGONKELEROKLENAKELGK 290  
 QY 219 IKRTIASKVSLGAVFPMNGTYGLAFWGTSLINGEPGTYTGLVLAFFSVIHSYCIG 278  
 Db 291 IKKASIANISMGIAPLLIYASALAFWYGTSLVTSKE--YTGAMATVFESILIGAFSVG 348  
 QY 279 AAVHFEFPAIRGAAPHFEOYIDKKPSIDNSTAGYKPESEGTVEFENVPNSPSRS 336  
 Db 349 QAAACIDAFANRGAAYVFDIIDNNPKIDSPSERGHKPDSTIKNEFDVHFSTPSRANVK 408  
 QY 339 IKILGGLNLRIKSGEVALVGLNGSGKSTVVOQLQRLYDPDGFIMVDENDIRALNVHRHYD 398  
 Db 409 IKILGGLNLKVGSGQVALVSGSGCKSTVQQLQRLYDPDGTINIDGQDIRNFVNYLRE 468  
 QY 399 RDHIGVAVSOEPLVFGTTSNNIKYGRDVTDEEMERARANAADTFMEFPNKFTLVGEGK 458  
 Db 469 REIGVAVSOEPLVFGTTSNNIKYGRDVTDEEMERARANAADTFMEFPNKFTLVGEGK 528  
 QY 459 KGAOMSGGOKORAIARALVRNPKIILDEATSAIDSEKSAVOALAEK 507  
 Db 529 RGAOLSGGOKORAIARALVRNPKIILDEATSAIDSEKSAVOALAEK 577

## RESULT 9

MDR2\_MOUSE STANDARD; PRT; 1276 AA.  
 ID MDR2\_MOUSE  
 AC P21440;

DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).  
 GN PGY2 OR PGY-2 OR MDR2.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RX MEDLINE=88302195; PubMed=3405218;  
 RA Gros P., Raymond M., Bell J., Housman D.;  
 RT "Cloning and characterization of a second member of the mouse mdr  
 RT gene family.";  
 RL Mol. Cell. Biol. 8:2770-2778(1988).  
 RP (12)  
 RP SEQUENCE OF 1-23 FROM N.A.  
 RC STRAIN-BALB/C;  
 RA Kirschner L.S., Horvitz S.B.;  
 RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS  
 CC NOT CAPABLE OF CONFERRING DRUG RESISTANCE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE  
 CC RELATED BUT DISTINCT CELLULAR GENES.  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J03398; AAA39516.1; -;  
 DR EMBL: M74151; AAA39515.1; -;  
 DR PIR: A30409; DVMS2.  
 DR HSSP: P13569; INBD.  
 DR MGD: MGI:97569; PGY2.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR InterPro: IPR001140; ABC\_transport\_tmem.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
 KW Multigene family.  
 FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 52 75 POTENTIAL.  
 FT TRANSMEM 119 139 POTENTIAL.  
 FT TRANSMEM 188 208 POTENTIAL.  
 FT TRANSMEM 215 235 POTENTIAL.  
 FT TRANSMEM 296 316 POTENTIAL.  
 FT TRANSMEM 325 345 POTENTIAL.  
 FT DOMAIN 346 707 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 708 728 POTENTIAL.  
 FT TRANSMEM 753 773 POTENTIAL.  
 FT TRANSMEM 829 849 POTENTIAL.  
 FT TRANSMEM 850 870 POTENTIAL.  
 FT TRANSMEM 933 953 POTENTIAL.  
 FT TRANSMEM 970 990 POTENTIAL.  
 FT DOMAIN 991 1276 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 426 433 ATP (BY SIMILARITY).  
 FT NP\_BIND 1066 1073 ATP (BY SIMILARITY).  
 FT REPEAT 1 635  
 FT REPEAT 636 1276  
 SQ SEQUENCE 1276 AA; 140332 MW; A6C38DDC2C118EEF CRC64;

Query Match 56.4%; Score 1463; DB 1; Length 1276;  
 Best Local Similarity 53.9%; Pred. No. 1,5e-91;  
 Matches 285; Conservative 103; Mismatches 111; Indels 30; Gaps 5;  
 QY 1 MIIIGLASLVNACPLPLMLVGLGSGDNLSICLVQNTYSF-----FRLLTL----- 47  
 DB 54 MFLGLTMAIHAHSGPLPMIVFGEMTKFDV-----NFGNSLPVNFSLMNPRIILE 107  
 QY 48 -----YVVGIGVAALIFGYIQLSWLITTAQRQRIKQPFHVSVALADIGWEDSCDI 99  
 DB 108 EEMTRAYVYSGIGGVAVIAVYQVSFWTLAAGRKIKRQKFHAILRQEMGFIDIKGT 167  
 QY 100 GELNTRMT-DIDKISDGIQDKIALLLFQNNSTSSIGLAVLGKMTLTVLTSPLIMAS 158  
 DB 168 TELNTRRLTDVSKISGIDCKVGMFQALATFAGFVGFIRGKTLTLMASPLTGLIS 227  
 QY 159 AAACSRMVISLTKSLSAVSKAGAAVEYLSSIRTVIAFRADKELORYTONLKADKDFG 218  
 DB 228 TAVMAKILSTFSDKELAAVAKAGAAVEAPGAIKTVIYAGGONKLEKRYOKHLENAKKIG 287  
 QY 219 IKRTASKVSLGAVYFFNNGTGLAFWGTSLINCEPGYITGVYAVFVSHSSYCIQ 278  
 DB 288 IKKASISANISMGIAFLIYASVALAWGSTLYISKE--YTGAMATVFFSILIGAFSVG 345  
 QY 279 AAVPHETFAIRGAFFHFOYIDKRPSTDNFSTAGIKRESIEGYEFKNVSPNYSRPS 338  
 DB 346 QAAPCIDAFANRGAAYVFIIDNNPKIDSPSEHKKDNNKGNLEFSDVHSYSRAN 405  
 QY 339 IITLGLNLRIRSGEVALVAVGLSGSTVVOLORLYPPDGFIVDENDIRALVRRY 398  
 DB 406 IITLGLNLRIRSGEVALVAVGLSGSTVVOLORLYPPDGFIVDENDIRALVRRY 465  
 QY 399 RDHIGVSOEPLVFGTTISNNIKYGRDVTDEEMERAAEANAAYDIMEPNKFNLTVE 458  
 DB 466 RELIGVSGEPLVFGTTISNNIKYGRDVTDEEMERAAEANAAYDIMEPNKFNLTVE 525  
 QY 459 KAGNSGGGOKORAIARALVRNPKIILDEATSAIDSEKSAVOALDK 507  
 DB 526 RGAQLSGGOKORAIARALVRNPKIILDEATSAIDSEKSAVOALDK 574  
 RESULT 10  
 MDR2\_RAT ID STANDARD; PRT; 1278 AA.  
 AC Q08201;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MULTIDRUG RESISTANCE PROTEIN 2 (P-GLYCOPROTEIN 2).  
 GN PGY2 OR MDR2.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RX Nucleic Acids Res. 21:3885-3891(1993).  
 RT "Cloning and regulation of the rat mdr2 gene."  
 RL Brown P.C., Thorgeirsson S.S., Silverman J.A.;  
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED  
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L15079; AAA02937.1; -  
 DR HSSP: P13569; 1MB.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR InterPro: IPR001140; ABC\_transportr\_tmem.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;  
 KW Multigene family.  
 FT DOMAIN 1 54  
 FT TRANSMEM 1 54 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 119 75 POTENTIAL.  
 FT TRANSMEM 188 139 POTENTIAL.  
 FT TRANSMEM 208 208 POTENTIAL.  
 FT TRANSMEM 215 235 POTENTIAL.  
 FT TRANSMEM 296 316 POTENTIAL.  
 FT TRANSMEM 325 345 POTENTIAL.  
 FT DOMAIN 346 709 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 710 730 POTENTIAL.  
 FT TRANSMEM 755 775 POTENTIAL.  
 FT TRANSMEM 831 851 POTENTIAL.  
 FT TRANSMEM 852 872 POTENTIAL.  
 FT TRANSMEM 935 955 POTENTIAL.  
 FT TRANSMEM 972 992 POTENTIAL.  
 FT DOMAIN 993 1278 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 426 433 ATP (POTENTIAL).  
 FT NP\_BIND 1068 1075 ATP (POTENTIAL).  
 SO SEQUENCE 1278 AA; 140655 MW; DAFB6BE745AF73BF CRC64;

Query Match 55.7%; Score 1445; DB 1; Length 1278;  
 Best Local Similarity 53.1%; Pred. No. 2,4e-90;  
 Matches 281; Conservative 105; Mismatches 113; Indels 30; Gaps 5;

1 MIIIGLILVNGACLPMPVILGEMSDNLISGCLVNTYSE-----FRLTL----- 47  
 54 MLUSTAMAHGSLPLMIVFGMTKPD-----NAGFSLPVNFSLMLPGRITL 107  
 48 -----YVVGIGVAALIFGYIQISLMIITARQTRIRKQFPHSVLAODIGFDS 99  
 108 EEMTRVAYVYSGGGLVLAIVGVSEFWTLAAGQIRIRKQFPHALIRQEMGFIDIGT 167  
 100 GELNTRMT-DIDKISDGIKRIALLFQNMSTFSGVLAVGKVKRLTYLTSTPLIMAS 158  
 168 TELNTRLTLDISKISEGIGDKVGFQAIATFAGFTVGRKMLTLVIMATITLIGLS 227  
 159 AAASRWVITSLTSELSKELSVKAGVAEVLSSIRTVIAFRAQEKELQRYTONLDAKDFG 218  
 228 TAYWAKLITSTFSDKELAYAKAGAAEALGAIKRVIAFGQNKLEKRYKHLEAKKIG 287  
 219 IKRTIRKSVLSGAVYFPMNGTYGLAFWYGTSLINGEPGYTIGVLAVFVSHSYCIG 278  
 288 IKKAIISANISMGIAFLIIVASYALAFWYGTSLVYSKE--YITGNAMTYVFSLIGAFSVG 345  
 279 AAVHFEFTAIARAGAFHIFQVYIDKPSIDNFSTAGRYKPESTIEGVVEKKNVSFNPSPS 338  
 346 QAAFCIDAFAPNARAAAYVIFDIIDNPIDTSERGHKPDISKGNLEFSDVHFVSBRAN 405  
 339 IKILKGNLIRKSGEYVALVINGSGKSTVOLLRLDPDGGFTMVDENDIRALNVHY 398  
 406 IKILGNLAKKSGGOTVALVINGSGKSTVOLLRLDPDGGFTMVDENDIRALNVHY 465  
 399 RDHIGVSGQEPVLEFQTTISNNIKYGRDVTDEMEARAAREANAYDFIMEFNKFTLVGE 458  
 466 REFIGVSGQEPVLEFQTTIAENIRIRGVNWTDEIKAVKEANAYDFIMKLPQKFTLVGD 525  
 459 KGAONGSGOKORAIARALVRNPKTLILDEATSLDSESKSAVQAALEK 507

DB 526 RGAOLSGGOKORAIARALVRNPKTLILDEATSLDSESKSAVQAALEK 574

RESULT 11  
 ID AB11\_HUMAN STANDARD; PRT; 1321 AA.  
 AC 095342; 09UNB2;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11).  
 GN ABCB11 OR BSEP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT PFIC2 GLY-297.  
 RX MEDLINE-99021377; PubMed-9806540;  
 RA Strautniece S.S., Bull L.N., Kocoshis S.A., Dahl N.,  
 RA Arneil H., Sokal E., Dahan K., Childs S., Ling V., Tanner M.S.,  
 RA Kagawa A.A., Nemeth A., Pawlowska J., Baker A., Mielit-Vergani G.,  
 RA Freimer N.B., Gardiner R.M., Thompson R.J.;  
 RA "A gene encoding a liver-specific ABC transporter is mutated in  
 RT progressive familial intrahepatic cholestasis.";  
 RL Nat. Genet. 20:233-238(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Mol O., Hooiveld G.J.E.J., Jansen P.L.M., Muller M.;  
 RT "Cellular localization and functional characterization of the human  
 RT bile salt export pump (BSEP).";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
 CC INTO THE CANALICULUS OF HEPATOCYTES.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
 CC SITU (BY SIMILARITY).  
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
 CC ATP BINDING CASSETTE (ABC) DOMAIN.  
 CC -1- DISEASE: DEFECTS IN ABCB11 ARE THE CAUSE OF PROGRESSIVE FAMILIAL  
 CC INTRAHEPATIC CHOLESTASIS 2 (PFIC2), AN INHERITED LIVER DISEASE OF  
 CC CHILDHOOD. PFIC2 IS CHARACTERIZED BY CHOLESTASIS AND NORMAL SERUM  
 CC GAMMA-GTUTAMYLTRANSFERASE ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).







```

FT DOMAIN 375 755 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 756 776 POTENTIAL.
FT DOMAIN 777 794 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 795 815 POTENTIAL.
FT DOMAIN 816 869 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 870 890 POTENTIAL.
FT TRANSMEM 891 911 POTENTIAL.
FT DOMAIN 912 979 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 980 1000 POTENTIAL.
FT TRANSMEM 1001 1011 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1012 1032 POTENTIAL.
FT DOMAIN 1033 1321 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 1455 1462 ATP (POTENTIAL).
FT NP_BIND 1113 1120 ATP (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1321 AA: 146257 MW: 54434E7B9FB1F6 CRC64:

```

Query Match 48.0%; Score 1243.5; DB 1; Length 1321;  
 Best Local Similarity 47.0%; Pred. No. 1.2e-76;  
 Matches 255; Conservative 98; Mismatches 153; Indels 37; Gaps 4;

```

OY 1 MIIIGLASLVNGACPLMPLVLGEMSDNLI-----SGCLVQNTY----- 40
DB 62 MIMGVALLHGMAGPQILIFGIMDFIKYDIEROELEIPKACVNNITVINSFHQ 121
OY 41 -----SFFRLTYVVGVAALLFGYQISLMTITARQRIKRIKOPFHS 85
DB 122 NMNTNGVCGLVDESEMIKFSIYAGVMTVLILGYFOILWITGRQIRRMKRIERR 181
OY 86 VLAODIGWEDSCDIGELNTWT-DIKISDGIQKIALLRQNMSTFSGIAGVKGKML 144
DB 182 IMMEIGWECTSYGELNSRFADDEKINDAIDQLAHFIQRMSTANGCILLGFRGKML 241
OY 145 TLVTLSTPLIMASAAACSRVLSLSKELSAVSKAGAVAEVLSIRTYIAFPAQEKEL 204
DB 242 TLVLAVSPILGIAAIVIGLSIAKFELEKAKAGSIDDEVLSIRTYIAFPGENKEV 301
OY 205 QRTONAKDKRQIKRTISKVSLSGAVYFPMNCTGLAFMYGSLILNGPGYTIGTVL 264
DB 302 ERKEKIVFQRMGIMGMWGEFTGIMCIIFFCYALAEVIGSTVLVD-EETTPGTLV 360
OY 265 AVFEVSIVHSYCIGAAPHETFAIARGAAFHIFQVIDKPSIDNFSTAGYKPSISGTV 324
DB 361 QIFLCVILAMNIGHASSCLEIFSTGSAANNIFQITDRQVIDCMGSDGYKIDRIRGEI 420
OY 325 EFKNSFVPSRPSIKILKIGLNRKSGEYVALVGLNGSKSVYVOLLQRLYDDDGFI 384
DB 421 EFHNVTFHYSRPRDKVILDNLSMVIKPGETALVSGSAGAKSTALLOLQRFYDCEGV 480
OY 385 VDEMDIALNVRHRIHGVYSEPVLEFQTISNNIKYGRDVADEMERARANAYDF 444
DB 481 LDGHDISLRIRWLRDGIYEDPEVLEFSTIAENIRREDATMEDIVQAKANANF 540
OY 445 IMEPKENTLVGEKGAGSGORAIARALRNPKIILDEATSAIDESSAVQAA 504
DB 541 IMALPQFDTLVGGGGGSGGQKORVAIARALRNPKIILDMATSLDNESERVOEA 600
OY 505 LEK 507
DB 601 LNK 603

```

```

DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)
GN (SISTER OF P-GLYCOPROTEIN).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=20076398; PubMed=10607905;
RT Green R.M., Hoda F., Ward K.L.;
RT "Molecular cloning and characterization of the murine bile salt export
RL pump.";
RL gene 241:117-123(2000).
RN [2]
RP SEQUENCE OF 463-635 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RA Salkar R., Suchy F.J., Ananthanarayanan M.;
RT "Molecular cloning of mouse liver bile salt export pump (bsep).";
RT Submitted (SEP-1999) to the EMBL/GenBank/DDP databases.
RL -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
CC INTO THE CANALICULUS OF HEPATOCYTES.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
CC SITU.
CC -1- DOMAIN: MULTIFUNCTIONAL. POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC ATP BINDING CASSETTE (ABC) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF133903; AFI14372.1; -.
DR EMBL: AF186585; AAD56419.1; -.
DR MGD: MGI:1351619; Abcb11.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR000439; ABC_transportr.
DR InterPro: IPR001140; ABC_transportr_tmern.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transmembrane; Transport.
FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 63 83 POTENTIAL.
FT DOMAIN 84 147 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 148 168 POTENTIAL.
FT DOMAIN 169 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 POTENTIAL.
FT DOMAIN 237 240 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 241 261 POTENTIAL.
FT DOMAIN 262 319 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 320 340 POTENTIAL.
FT DOMAIN 341 353 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 354 374 POTENTIAL.
FT DOMAIN 375 755 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 756 776 POTENTIAL.
FT DOMAIN 777 794 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 795 815 POTENTIAL.
FT DOMAIN 816 869 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 870 890 POTENTIAL.
FT TRANSMEM 891 911 POTENTIAL.

```

FT DOMAIN 912 979 CYTOPLASMIC (POTENTIAL).  
FT TRANSSEM 980 1000 POTENTIAL.  
FT DOMAIN 1001 1011 EXTRACELLULAR (POTENTIAL).  
FT TRANSSEM 1012 1032 POTENTIAL.  
FT DOMAIN 1033 1321 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 455 462 ATP (POTENTIAL).  
FT NP\_BIND 1113 1120 ATP (POTENTIAL).  
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 481 481 L -> P (IN REF. 2).  
FT CONFLICT 633 633 T -> V (IN REF. 2).  
SO SEQUENCE 1321 AA; 146675 MW; 1585EBF175D32967 CRC64;

Query Match 47.4%; Score 1227.5; DB 1; Length 1321;  
Best Local Similarity 47.1%; Pred. No. 1.5e-75;  
Matches 256; Conservative 97; Mismatches 155; Indels 39; Gaps 6;

OY 1 MILGLASLVNGACLPMLPVLGEMSD----- 27  
D 62 MEMGSVCLALGMAQPGMIYVFGILTDIFVEYDIEROELSPGKVCMMNTIWINSSFNO 121  
OY 28 NLISG-C-LVQNTNYSFRLTYVGVVALIFGYIOISLWITTAAROTKRIKQFPH 84  
D 122 NMTNSTGGLVDINS-EVIRKSGIYAGVAVLLIGYFQIRLWITVGARQIKRMREYFR 180  
OY 85 SVLADIGMFPSCDIGELINTRT-DIDKISDIGNKIALLFONMSTFISGLAVGLVGMK 143  
D 181 RIMREIGMFDCTSVGELNSRSDINKIDEALDQMALFLQRLSTALSGLLIGFYRGK 240  
OY 144 LTLVTLSTPLIMASAAACSRNIVISLTSKELSAKAGAVAEVLSITFYIAFRAQKE 203  
D 241 LTLVTLAVSPLIGIAVAIGLSVAKFTELEKAKAGSIDAEVLSITFYIAFGGENKE 300  
OY 204 LORYONLKDKEGFKRTIASKVSGLGAVYFPMNGTYGIAFYGSLTINGEPYTGIV 263  
D 301 VERYKKNLMPQRMKIMGMVWGFTGYWCLIFCYALAFYGSRLVID-EGEYTPGL 359  
OY 264 LAVFESVHSYICGAAPVHETFAIARGAFHIFQVIDKKISINFSTAGKPSIEGT 323  
D 360 IQIFLCVLIAMANNIGNASSCLEIFSTGCSAASSIFQIDRQVCMQMSDQKILDRIKE 419  
OY 334 VEKRVSNFSPRSIKILKGLINIKSGEVALVGLNGSGKSTVOLLQRLYDPDGT 383  
D 420 IEFHNVTFHYSRPEVKILNNLSMVIKPGETAFFVGSSGAKSTALQLQRFRECEGW 479  
OY 384 MVDEMDIALNVRHRYRDHGVVSOEPLFGTTISNNIKYGRDQVDEMERAREANAYD 443  
D 480 TLDGHDLSLNRDLDOIGIVEDEPLFSTTIAENIRLGRREATMEDTVOAKADNAYN 539  
OY 444 FIMEPKNFNTLVGEKGAQMSGQKORIAIARALVRNPKILLDEATSAIDSEKSAVOA 503  
D 540 FIMALPQGFDTLVGGGGGSGQKORAVALARALKRPRILLDMATSLADNESEAKYVG 599  
OY 504 ALEK 507  
D 600 ALNK 603

RESULT 14  
ID AB11\_RABIT STANDARD: PRT: 1321 AA.  
AC Q9NOV3:  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE BILE SALT EXPORT PUMP (ATP-BINDING CASSETTE, SUB-FAMILY B, MEMBER 11)  
DE (SISTER OF P-GLYCOPROTEIN).  
GN ABCB11 OR BSEP OR SPGP.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NEW ZEALAND WHITE; TISSUE=Liver;  
RA Balasubramanian N.V., Suchy F.J., Ananthanarayanan M.;  
RT "Molecular cloning and characterization of rabbit liver bile salt  
export pump (Bsep/spgp).";  
RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.  
CC -I- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS  
CC INTO THE CANALICULUM OF HEPATOCYTES.  
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN  
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR  
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN  
CC SITU.  
CC -I- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
CC ATP BINDING CASSETTE (ABC) DOMAIN.  
CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sb.ch/announce/>  
CC or send an email to [license@isb-sb.ch](mailto:license@isb-sb.ch)).  
CC -----  
CC EMBL: AF249879; AAF65552.1; -.  
DR InterPro: IPR003593; AAA, transporth.  
DR InterPro: IPR003439; ABC\_transport.  
DR InterPro: IPR001140; ABC\_transport\_tmemb.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane\_2.  
DR Pfam: PF00005; ABC\_tran; 2.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
KW ATP-binding; Transmembrane; Transport.  
FT DOMAIN 1 62  
FT TRANSSEM 63 83  
FT DOMAIN 84 147  
FT TRANSSEM 148 168  
FT DOMAIN 169 215  
FT TRANSSEM 216 236  
FT DOMAIN 237 240  
FT TRANSSEM 241 261  
FT DOMAIN 262 319  
FT TRANSSEM 320 340  
FT DOMAIN 341 353  
FT TRANSSEM 354 374  
FT DOMAIN 375 755  
FT TRANSSEM 756 776  
FT DOMAIN 777 794  
FT TRANSSEM 795 815  
FT DOMAIN 816 869  
FT TRANSSEM 870 890  
FT DOMAIN 891 911  
FT TRANSSEM 912 979  
FT DOMAIN 980 1000  
FT TRANSSEM 1001 1011  
FT DOMAIN 1012 1032  
FT TRANSSEM 1033 1321  
FT NP\_BIND 455 462  
FT NP\_BIND 1113 1120  
FT CARBOHYD 109 109  
FT CARBOHYD 116 116  
FT CARBOHYD 122 122  
FT CARBOHYD 125 125  
SO SEQUENCE 1321 AA; 146376 MW; 457539FCD6D717A2 CRC64;

[illegible][illegible]

```

OY 328 NVSFNYPSPRSIKILKGINLRIKSGEYVALVGLNGSGKSTVVQLLQRLYDDPDGFIMWE 387
Db 406 GIRFRTPARPVEILKGLTVDPGLPGQYVAFVGASGCGKSTLIQLMQRFYDPEAGSVKLDG 465
OY 388 NDIRALNVRYRDHIGVYVSOEPVLEFQTTISNNIKYGRDDVTDEEMERAAAREANAYDFIME 447
Db 466 RDLRTLVNGWLRSGQICGVVGQEPVLEFATTIGENIRYGRPSATQADI EKARANANCHDFITR 525
OY 448 FPNKENTLVGEKGQOMSGGOKORIAIARALVBNPKILILDEATSAIDSEKSAVOALE 506
Db 526 LPKGIDYVGEKGQISGGOKORIAIARALVROPQVLLDEATSAIDPTSEKVOALE 584

```

Search completed: April 22, 2002, 14:37:51  
 Job time: 303 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:37:10 ; Search time 76.9 Seconds

(without alignments)  
977.685 Million cell updates/sec

Title: US-09-873-409-8  
Perfect score: 2592  
Sequence: 1 MILGILASLVNGACLEPLMPL.....SEKSAVQALAEKDPYRSF 514

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_17:\*  
2: SP\_archaea:\*  
3: SP\_bacteria:\*  
4: SP\_fungi:\*  
5: SP\_human:\*  
6: SP\_invertebrate:\*  
7: SP\_mammal:\*  
8: SP\_mhc:\*  
9: SP\_organelle:\*  
10: SP\_phage:\*  
11: SP\_plant:\*  
12: SP\_rodent:\*  
13: SP\_virus:\*  
14: SP\_vertebrate:\*  
SP\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2346.5	90.5	469	4	014573 homo sapien
2	1576	60.8	1280	6	046605 canis famill
3	1571.5	60.6	1272	11	09JK64 O9tKu6 ratius norv
4	1563.5	60.3	1163	6	09TSU2 felis silve
5	1548.5	59.7	1285	5	002793 ovis aries
6	1546.5	59.7	1287	13	091586 xenopus lae
7	1508.5	58.2	1287	13	093437 gallus gall
8	1497	57.8	1169	11	060502 cricetus
9	1243.5	48.0	1321	11	088331 ratius norv
10	1158	44.7	1294	5	045721 caenorhabd
11	1081.5	41.7	1289	5	001495 caenorhabd
12	1065	41.1	1313	5	09V626 drosophila
13	1050.5	40.5	1275	5	061301 haemochus
14	1033	39.9	1285	10	09L6X1 O9tKu6 ratius norv
15	1025.5	39.6	1279	5	09V616 drosophila
16	1023	39.5	1321	5	021349 caenorhabd
17	1017	39.2	1286	10	080725 arabidopsis
18	1008	38.9	1230	10	09SVT3 O9tKu6 ratius norv
19	1003.5	38.7	1313	10	09ZRG2 solanum tub

20	1000.5	38.6	1283	5	024393 drosophila
21	998	38.5	1292	10	09M109 O9tKu6 ratius norv
22	996	38.4	1229	10	09SVT2 O9tKu6 ratius norv
23	993	38.3	1229	10	049749 arabidopsis
24	986.5	38.1	1248	10	09FHF1 O9tKu6 ratius norv
25	985.5	38.0	1222	10	09SVW6 O9tKu6 ratius norv
26	985.5	38.0	1323	10	09MOM2 O9tKu6 ratius norv
27	985	38.0	1278	10	09FMX7 O9tKu6 ratius norv
28	977.5	37.7	1318	5	019733 caenorhabd
29	975.5	37.6	1408	3	043140 cryptococcu
30	975	37.6	1348	3	09Y8G1 emerlicella
31	972.5	37.5	1268	5	020335 caenorhabd
32	971.5	37.5	1266	5	020331 caenorhabd
33	966	37.3	1229	10	09FMX8 O9tKu6 ratius norv
34	966	37.3	1292	3	09C163 rhizomucor
35	965.5	37.2	1286	10	09ZRR2 O9tKu6 ratius norv
36	959	37.0	1349	3	043121 aspergillus
37	957.5	36.9	1341	5	09U7F4 onchocerca
38	956	36.9	1341	5	015903 leishmania
39	950.5	36.7	1302	5	024851 entamoeba h
40	945.5	36.5	1217	3	09HGT5 O9tKu6 ratius norv
41	944.5	36.4	1310	5	024852 entamoeba h
42	940.5	36.3	1245	10	09C7F8 O9tKu6 ratius norv
43	939	36.2	1341	5	096844 leishmania
44	931	35.9	1341	5	09XXU8 leishmania
45	927.5	35.8	1252	10	09LJX0 O9tKu6 ratius norv

## ALIGNMENTS

RESULT 1  
ID 014573 PRELIMINARY; PRT; 469 AA.  
AC 014573;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE WUGSC:H\_RG367017.1 PROTEIN (FRAGMENT).  
GN WUGSC:H\_RG367017.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kalicki J., Smith A., Gibson A.;  
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Waterston R.;  
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).  
DR EMBL: AC002486; AAB67603.1; -  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 1.  
DR Pfam: PF00005; ABC\_tran; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
KW ATP-binding; Transport.  
FT NON\_TER 1 469  
FT TER 469  
SQ SEQUENCE 469 AA; 51667 MW; E67E088FF594FC3 CRC64;

Query Match 90.5%; Score 2346.5; DB 4; Length 469;  
Best Local Similarity 99.8%; Pred. No. 2.5e-148;  
Matches 468; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 40 YSFRTLYVVGVALIFGYIOISLMTITARQTRIKRPFHSLVLAODIGWPFSCDI 99  
|||||

```

Db 1 YSFRLTYVIGVAAIFGIQISLMIITTAQOTKRIKOPFHSHVLAODIGWFGSDI 60
QY 100 GELNTRMT-DIDKISDGDGKIALLFQNMSTFSGIAGLVKGKMLTYLTSTPLIMAS 158
Db 61 GELNTRMTDIDKISDGDGKIALLFQNMSTFSGIAGLVKGKMLTYLTSTPLIMAS 120
QY 159 AAACSRWVISTLSKELSAYSKAGAAVEEVLSSIRTVIAFRAQEKELQRYTONLKDADFG 218
Db 121 AAACSRWVISTLSKELSAYSKAGAAVEEVLSSIRTVIAFRAQEKELQRYTONLKDADFG 180
QY 219 IKRTIAKSVSLGAVYFFPMNGTYGLAFWYGTSLINGEPGYTIGYLAFFSVIHSSYICG 278
Db 181 IKRTIAKSVSLGAVYFFPMNGTYGLAFWYGTSLINGEPGYTIGYLAFFSVIHSSYICG 240
QY 279 AAVHFEFATARGAAFIIFQVIDKPSIDNFSTAGYKPESEIEGVEKKNVSPNPSRPS 338
Db 241 AAVHFEFATARGAAFIIFQVIDKPSIDNFSTAGYKPESEIEGVEKKNVSPNPSRPS 300
QY 339 IKILGLNLRIKSGEYVALVGLNGSGKSTVYQLQRLYPDPDFIMVENDIRALNVRY 398
Db 301 IKILGLNLRIKSGEYVALVGLNGSGKSTVYQLQRLYPDPDFIMVENDIRALNVRY 360
QY 399 RDHIGVVSQEPVLFGTITSNRIKGRDVTDEEMERARAENAYDFIMEFNKENTLVE 458
Db 361 RDHIGVVSQEPVLFGTITSNRIKGRDVTDEEMERARAENAYDFIMEFNKENTLVE 420
QY 459 KGAOMSGGOKORIAIARALVNPRLILDEATSLDSEKSAVOALEK 507
Db 421 KGAOMSGGOKORIAIARALVNPRLILDEATSLDSEKSAVOALEK 469

```

```

RESULT 2
ID 046605 PRELIMINARY; PRT: 1280 AA.
AC 046605:
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MULTIDRUG RESISTANCE P-GLYCOPROTEIN.
OS MDRL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON.
RA Puel O., Lepage J.F., Alvine M., Galtier P., Pineau T.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL: AF045016; AAC02113.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1280 AA; 141523 MW; 762D5AFA7C73306 CRC64;

```

```

Query Match 60.8%; Score 1576; DB 6; Length 1280;
Best Local Similarity 57.1%; Pred. No. 2.5e-96;
Matches 302; Conservative 98; Mismatches 105; Indels 24; Gaps 4;
QY 1 MIIIGTILASVNGACLPIMPLVIGEMSDNLISGCL-----VQNTYSFFR--- 44
Db 50 MIVGTMAIIHGALPLMLPLFGNMTSEFANAGISRNKTFPVIIINESTNTNQHINLE 109
QY 45 -----LTLVYVIGVAAIFGIQISLMIITTAQOTKRIKOPFHSHVLAODIGWFGSDI 99

```

```

Db 110 EEMTYAYVYSGIGAGVLAAYIOVSWFCCLAAGROILKIRROFPAHMRQELIGMDVADV 169
QY 100 GELNTRMT-DIDKISDGDGKIALLFQNMSTFSGIAGLVKGKMLTYLTSTPLIMAS 158
Db 170 GELNTRMTDIDKISDGDGKIALLFQNMSTFSGIAGLVKGKMLTYLTSTPLIMAS 120
QY 159 AAACSRWVISTLSKELSAYSKAGAAVEEVLSSIRTVIAFRAQEKELQRYTONLKDADFG 218
Db 121 AAACSRWVISTLSKELSAYSKAGAAVEEVLSSIRTVIAFRAQEKELQRYTONLKDADFG 180
QY 219 IKRTIAKSVSLGAVYFFPMNGTYGLAFWYGTSLINGEPGYTIGYLAFFSVIHSSYICG 278
Db 181 IKRTIAKSVSLGAVYFFPMNGTYGLAFWYGTSLINGEPGYTIGYLAFFSVIHSSYICG 240
QY 279 AAVHFEFATARGAAFIIFQVIDKPSIDNFSTAGYKPESEIEGVEKKNVSPNPSRPS 338
Db 241 AAVHFEFATARGAAFIIFQVIDKPSIDNFSTAGYKPESEIEGVEKKNVSPNPSRPS 300
QY 339 IKILGLNLRIKSGEYVALVGLNGSGKSTVYQLQRLYPDPDFIMVENDIRALNVRY 398
Db 301 IKILGLNLRIKSGEYVALVGLNGSGKSTVYQLQRLYPDPDFIMVENDIRALNVRY 360
QY 399 RDHIGVVSQEPVLFGTITSNRIKGRDVTDEEMERARAENAYDFIMEFNKENTLVE 458
Db 361 RDHIGVVSQEPVLFGTITSNRIKGRDVTDEEMERARAENAYDFIMEFNKENTLVE 420
QY 459 KGAOMSGGOKORIAIARALVNPRLILDEATSLDSEKSAVOALEK 507
Db 421 KGAOMSGGOKORIAIARALVNPRLILDEATSLDSEKSAVOALEK 469

```

```

RESULT 3
ID 09JK64 PRELIMINARY; PRT: 1272 AA.
AC 09JK64:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN 1A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=LIVER;
RA Hooiveld G.J.E.J., Wilms J.W.J., Hagenbuch B., Jansen P.L.M.,
RA Meijer D.K.F., Muller M.;
RT "Cloning and functional characterization of the rat multidrug
RT resistance protein Mdr1a.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL: AF257746; AF69007.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1272 AA; 140326 MW; B5C1DE89DFC986C5 CRC64;

```

```

Query Match 60.6%; Score 1571.5; DB 11; Length 1272;
Best Local Similarity 57.7%; Pred. No. 4.9e-96;
Matches 301; Conservative 101; Mismatches 101; Indels 19; Gaps 4;
QY 1 MIIIGTILASVNGACLPIMPLVIGEMSDNLISGCLVQNTYSFFRL----- 46

```

```
Db 50 MLGLTALIHGIALPLMLVFGDMTFDFAN--VGNNSMSMYNNTDIAKLEDEMTTYA 107
Oy 47 LYVVGIGVAAALIFGYIOISLMTITTAARQTKRIKOFHSHVLAQDIGWPDSCDIGELNTRM 106
Db 108 YYTIGIGAGVILVAVIOWSLMCLAGROIHKIRQKFFHAIMQOISGIMPDVHDVGEINTRL 167
Oy 107 T-DIDKIDSGDKTALLFQNMSTFISGLAVGVKMKLTIVLTSLTPILMASAACSRLM 165
Db 168 TDDVSKINEGIGDKIGMFQAMATFFGGFIIGFRGKMLTIVLILASIPVLGISAGIMAKI 227
Oy 166 VISLTKSLTASYKAGAAVEVLSIRTVIAFRAQEKELORYTONUKAKDGIKRTIAS 225
Db 228 LSTFTDKLOAVAKAGAAVEVLAIRTVIAFGGKKELERYNNMLEAKRIGIKKAITA 287
Oy 226 KVSIGAVYEFNNGYGLAFWYGTSLILNGEPYTGIVLAVFFSVIHSYICGAAPHRE 285
Db 288 NISMAAEFLLIYASVALAFWYGTSLIVLSKE--YTIQVLTFFSVLIGAFSVGQASPNI 345
Oy 286 TFLATARGAAFIHFOYIDKKPSIDNFTAGYKPESTIEGTVFERNNVSPNRPSTIKILGL 345
Db 346 AFANARGAAVEVFSIIDMKPSIDFSKSGHKPDNIQNLFEKNIHFYSYPSRKDVOLKGL 405
Oy 346 NLRIKSGEVALVGLNGSGKSTVVOQLRLYDPDGFIMVDENDIRALNVRHYRDIHGVY 405
Db 406 NLKRSQGTVALVNGSGCKSTVOLLRLDPIGCVSISDQDTRTINVRILREILIGVY 465
Oy 406 SOEPVLTGTTISNNIKYGRDVTDEEMERARANAYDFIMEPPNKFNTLVGKQAGMSG 465
Db 466 SOEPFLATTAENIRYGRENVYTMDEIKAVKEANAYDFIMKLPKFPILVGERGAOLSG 525
Oy 466 GOKORIALRALVRRPKLILIDEATSLDSEKSAVOALEK 507
Db 526 GOKORIALRALVRRPKLILIDEATSLDSEAVVOALDK 567

RESULT 4
O9TSU2 PRELIMINARY; PRT: 1163 AA.
AC O9TSU2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MUTTI-DRUG RESISTANCE RELATED MRNA, PARTIAL CDS (FRAGMENT).
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Okai Y., Nakamura N., Matsusito H., Kato H., Setoguchi A., Yazawa M.,
RA Okuda M., Wataai T., Hasegawa A., Tsujimoto H.;
RT "Molecular analysis of multidrug resistance in lymphoma cells in the
RT cat.";
RL Am. J. Vet. Res. 0-0-0(1999).
CC -!- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL; AB029153; BA87071.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 2.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
FT NON_TER 1
FT NON_TER 1163
SQ SEQUENCE 1163 AA; 128510 MW; 1B5B413776A93A26 CRC64;
```

Query Match 60.3%, Score 1563.5; DB 6; Length 1163;

```
Best Local Similarity 58.5%; Pred. No. 1,4e-95;
Matches 303; Conservative 93; Mismatches 99; Indels 23; Gaps 4;

Oy 11 NCACLPMLPVLIGMSDNL-----ISGLVQNTYSPFR-----TLVVV 50
Db 1 HCAALPLMLVFGDMTFDFANAGISRNLTTLTNGESIYNDSYFIRLEEMTVAAYVS 60
Oy 51 GIGVAAALIFGYIOISLMTITTAARQTKRIKOFHSHVLAQDIGWPDSCDIGELNTRT-DI 109
Db 61 GIGAVLVAAAYIQVSWFCMLAAGROILKIRKOPFHAIMREQVMPFVDHVGELNTRLTDV 120
Oy 110 DISIDGDKTALLFQNMSTFISGLAVGVKMKLTIVLTSLTPILMASAACSRLM 169
Db 121 SKINIEGDKTIGMFQAMATFFGGFIIGFRGKMLTIVLILASIPVLGISALIAKLSIF 180
Oy 170 TSKELASVSKAGAAVEVLSIRTVIAFRAQEKELORYTONUKAKDGIKRTIASVSL 229
Db 181 TDKELLAAVAKAGAAVEVLAIRTVIAFGGKKELERYNNMLEAKRIGIKKAITANISI 240
Oy 230 GAVYEFNNGYGLAFWYGTSLILNGEPYTGIVLAVFFSVIHSYICGAAPHRETFAI 289
Db 241 GIAFLLIYASVALAFWYGTSLIVLSHE--YSIQVLTFFSVLIGAFSVGQASPSTIEAFAN 298
Oy 290 ARGAAFIHFOYIDKKPSIDNFTAGYKPESTIEGTVFERNNVSPNRPSTIKILGLNTRI 349
Db 299 AARGAAVEIFKIIDNKPSIDFSKNGHKPDNIKGNLEFKNVHSYPSRKEVLIKILKLNKY 358
Oy 350 KSGETVALVGLNGSGKSTVVOQLRLYDPDGFIMVDENDIRALNVRHYRDIHGVSOEP 409
Db 359 QSGQVVALVNGSGCKSTVOLLRLDPIGCVSISDQDTRTINVRILREILIGVSOEP 418
Oy 410 VIFGTTISNNIKYGRDVTDEEMERARANAYDFIMEPPNKFNTLVGKQAGMSGGOK 469
Db 419 VIFATTAENIRYGRENVYTMDEIKAVKEANAYDFIMKLPKFPDLYVGERGAOLSGGOK 478
Oy 470 RIAIRALVRRPKLILIDEATSLDSEKSAVOALEK 507
Db 479 RIAIRALVRRPKLILIDEATSLDSEAVVOALDK 516

RESULT 5
O02793 PRELIMINARY; PRT: 1285 AA.
AC O02793;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MULTIDRUG RESISTANCE PROTEIN-1.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Longley M., Crawford A.M.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC
CC TRANSPORTERS).
DR EMBL; U78609; AA858489.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Transport.
SQ SEQUENCE 1285 AA; 142020 MW; 90153E617C44856F CRC64;
```



Query Match 59.7%; Score 1548.5; DB 6; Length 1285;  
Best Local Similarity 56.3%; Pred. No. 1.7e-94;  
Matches 300; Conservative 100; Mismatches 100; Indels 33; Gaps 5;

OY 1 MILGLIASLVNACPLPMPVLVIGEMSDNLS-----GCLVQTN----- 38  
DB 55 MVLGLIAIIRHAGPLMLMLVFGDMTDFAGAGNIGTLISNTSTIDREYKLEK 114  
OY 39 ---TTSFRLTYVYGVAAALIFGYIOISLWITTAAROTKRIKOFHSHVLAODIGMPD 95  
DB 115 EMTTAY-----YSGIGAGVLIAYIOVSFMCCLAAGRVHRIKQFFHAIHQELEGWD 169  
OY 96 SCDIGELNTRMT-DIDKISDGIKIALLFQNMSTFSGIAGLVKGMKLTLYTSTSP 154  
DB 170 VHDVCKLNTLNDVSKINEGIGDKIGMFFQAMATFFETGIIIGTGMULTVILAIIV 229  
OY 155 IMASAACGRMYISLTSKLSAYKAGAAVEVLISRTVIAFRAQEKELORYTONLKA 214  
DB 230 IGLSAIYMKILISFDDKLLAYAKGAAVEVLAKIVIAFGGOKELERYKNLEEA 289  
OY 215 KDFGKRTASKVSLGAVYFFMNGTGLAFWYGTSLILNGEPYITIGTVLAVFVYIHS 274  
DB 290 KKGIGKATTAIINSMKAAALLIYASALAFWGTSLVLSRE--YSIGVLTVEFVSLIG 347  
OY 275 YCIGAAPHETFAIARGAAPHIIOVIDKKPSIDNFTAGYKEPSEIGVEFKVNSFNP 334  
DB 348 FSIQASPNIEAFANRGAAYEVEFKITDNKPSIDSYNTGCHKPDNIKGNLEFNVHHP 407  
OY 335 SRPSITLGLNLRISGFTVALVGLNSGKSTVOLLRLYDPDDGFIIVDENDIRALN 394  
DB 408 SRREVILGLNLRKVGSGGTVALVNSGCKSTVQLMOKLDPGTBMVSIDQDRIKTIN 467  
OY 395 VRRHYRIGVSOEPLFTGTTSNNIKYGRDVTDEEMERAREANAYDFIMEFPNKENT 454  
DB 468 VRLREITIGVSEPELFTTIAENIRYGEDVTYMEIQGVAEANAYPIMLPNKFDT 527  
OY 455 LVGEKAQMSGGOKRIAIARALVNRPKIILDEATSAIDSEKSAVOALEK 507  
DB 528 LVGERGAQLSGGOKRIAIARALVNRPKIILDEATSAIDSEKSAVOALEK 580

RESULT 6  
ID 091586 PRELIMINARY; PRT; 1287 AA.  
AC 091586;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE MULTIDRUG RESISTANCE PROTEIN.  
GN XEMDR.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8335;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95322451; Pubmed=7599185;  
RA Castillo G., Shen H.J., Horwitz S.B.;  
RT "A homologue of the mammalian multidrug resistance gene (mdr) is functionally expressed in the intestine of Xenopus laevis.";  
RL Blochim. Biophys. Acta 1262:113-123(1995).  
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).  
DB EMBL: U17608; AAA75000.1; -;  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003439; ABC\_transporter\_tmemb.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane\_2.  
DR Pfam: PF00005; ABC\_tran\_2.  
DR SMART: SM00382; AAA\_2.  
DR PROSITE: PS00211; ABC\_TRANSPORTER\_2.

KM ATP-binding: Transport.  
SQ SEQUENCE 1287 AA; 141504 MW; 06B95073C5771A15 CRC64;  
Matches 300; Conservative 91; Mismatches 114; Indels 17; Gaps 3;

OY 1 MILGLIASLVNACPLPMPVLVIGEMSDNLSGCLVQTNMTSFFRL-----T 46  
DB 66 MLFGTIAIASHAGALPLMLMLVFGEMTDSFVNGQVDTGNTWMSMINASRELQGMYYA 125  
OY 47 LYVGVGVALIFGYIOISLWITTAAROTKRIKOFHSHVLAODIGMPDSCDIGELNTR 106  
DB 126 YVYSGVGVVLCAYIIOISFWTLISAGROIKINSNFFHVLROEIGWFDINDGELNTRL 185  
OY 107 T-DIDKISDGIKIALLFQNMSTFSGIAGLVKGMKLTLYTSTSPILMASAACSRM 165  
DB 186 TDDVSKINEGIGKIMMLQSLTLTVGFTIGKGMKLTWVGALSPIMGLSAIYAKV 245  
OY 166 VISLTSKLSAYKAGAAVEVLISRTVIAFRAQEKELORYTONLKAADFGIKRTIAS 225  
DB 246 LSATFKKELKAYAKAGAAVEVLISRTVIAFGQNKELHRYERKNLEDAKKIGIKAITA 305  
OY 226 KVSIGAVYFFMNGTGLAFWYGTSLILNGEPYITIGTVLAVFVYIHSYCIAGAVPHE 285  
DB 306 NVISGAFMLTYAAYSLAFWYGTSLIIDG--GYTIGSVLTVEFVAVITIGAVGOTSPNIE 363  
OY 286 TFAIARGAAPHIIOVIDKKPSIDNFTAGYKPESEIGVEFKVNSFNPSPSITLIKGL 345  
DB 364 AFANARCAATITINIIDNGKIDISFSEKGLPKDKIKDIEFKNIFFYPSRKDIQVYKGL 423  
OY 346 NLRIKSGEVALVGLNSGKSTVOLLRLYDPDDGFIIVDENDIRALNRRHNDHIGVY 405  
DB 424 NLNPSGFTVALVSSGCKSTVOLLQREFYDEADVITLDDGDIRSLNIRYREITIGVY 483  
OY 406 SOEPVLFGTITSNNIKYGRDVTDEEMERAREANAYDFIMEFPNKENTLVGEKAQMSG 465  
DB 484 SOEPILEDITTIADIRKRGREYKKEELERATKEANADFIIMLPKLETLVGERGTLSG 543  
OY 466 GOKORIAARALVNRPKIILDEATSAIDSEKSAVOALEK 507  
DB 544 GOKORIAARALVNRPKIILDEATSAIDSEKSAVOALEK 585

RESULT 7  
ID 093437 PRELIMINARY; PRT; 1287 AA.  
AC 093437;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE ABC TRANSPORTER PROTEIN.  
GN CMDRL.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INTESTINE;  
RA Edelman H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C., Kane S.E., Kuchler K.;  
RT "Cmdrl, a chicken P-glycoprotein, confers multidrug resistance and mediates estrogen transport.";  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS).  
DB EMBL: AJ009799; CAA08835.1; -;  
DR HSSP: P13569; INBD.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmemb.

DR InterPro: IPR003439; ABC transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR InterPro: IPR000504; RRM.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR PROSITE: PS00300; RRM\_RNP\_1; UNKNOWN\_1.  
 DR ATP-binding; Transport.  
 KW SEQUENCE 1287 AA; 141785 MW; 3CFDBA8DD657111F CRC64;  
 SQ

Query Match 58.2%; Score 1508.5; DB 13; Length 1287;  
 Best Local Similarity 55.7%; Pred. No. 7.8e-92;  
 Matches 292; Conservative 95; Mismatches 118; Indels 19; Gaps 4;

QY 1 MILGILASLVNGACLPPLPLVVGEMSDNLI-----SGCLVQNTYTS-----FPR 44  
 DB 61 MIFGSLLAHAGTSLPIAMITFGDMTDSFVTCGMNTITGNSGLMSADVFKNLEEMR 120  
 QY 45 LTLVYGVGVALIFGYIQISLWITTAARQTKRIRKQEFHVSVAADIGMFDSCDIGELMT 104  
 DB 121 YAYYSAAIAAVLVAAVYIQTSEFTLAGROVKKIREKFEHAIMROEIGMFDVNDAGELMT 180  
 QY 105 RM-TDIDKISDGIKIALLPQNMSTFSGIAGLVKGGKLLVLTSTPLMASAACS 163  
 DB 181 RLIDVSKINSGIGDKIGFLIQSETTFLGTGFTGKRLTLVLAASPVGLSALMA 240  
 QY 164 RMVSLTSKELSAVSKAGAAVEEVLSTRTVAFRAQEKELORYONLMDAKDFGIRTI 223  
 DB 241 KILTFATKEDQAAVAKAGAAVEEVLAVRTVAFGQEKELIRYKHNLEDAKIRIGIRKAI 300  
 QY 224 ASKVELGAVFFPMNGTYGLAFYGTSLINGEPYTTIGTVLAVFVSIIHSSYICGAVPH 283  
 DB 301 TSNIMGAAFLIYASVIAFWYGTLLLANE--YSIGNVLVFFSVLLIGAFSIGOTAPS 358  
 QY 284 FETFAIANGAAPHIFQVIDKPKPSIDNEFSTAGKPSIEGTEFKNVSNFVPSRPIKILK 343  
 DB 359 IEAFNANAGAAVAFNIIDNEPEIDSYDAGHKPDHKKNLDFONFNFPSRPADELTK 418  
 QY 344 GLNLRKSGEYVALVGLNGSGKSTVQVLQRLYDPDGFIVNDENDIRALNRYHRDHIG 403  
 DB 419 GLNLKVGQGVYALVGGSCGKSTVQVLQRFYDPREGTITIDGDLKLANRYLREIIG 478  
 QY 404 VVSQEPVLFGTITSNKIKYGRDVTDEMERARAEANAYDFIMEPKNKNTLVGEKGAOM 463  
 DB 479 VYNQEPVLFATTAENIRYGRDVTMEERATKEANAYDFIMKLPKKEFVGEKGAOM 538  
 QY 464 SGGOKRIATARALVRNPKILLDEATSLDSEKSAVOALEK 507  
 DB 539 SGGOKRIATARALVRNPKILLDEATSLDSEKSAVOALEK 582

RESULT 8  
 060502 PRELIMINARY: PRT: 1169 AA.  
 AC 060502:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE P-GLYCOPROTEIN.  
 GN PGP-1.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OX NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91154265; PubMed=1671863;  
 RA Devine S.E., Hussain A., Davide J.P., Melera P.W.;  
 RT "Full length and alternatively spliced pgpl transcripts in multidrug-

RT resistant Chinese hamster lung cells.";  
 RL J. Biol. Chem. 266:4545-4555(1991).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 CC TRANSPORTERS).  
 DR EMBL: M59254; AAA37005.1; .  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmam.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR ATP-binding; Transport.  
 KW SEQUENCE 1169 AA; 128938 MW; 72E25B7CE29DC185 CRC64;  
 SQ

Query Match 57.8%; Score 1497; DB 11; Length 1169;  
 Best Local Similarity 60.4%; Pred. No. 3.9e-91;  
 Matches 284; Conservative 90; Mismatches 88; Indels 8; Gaps 3;

QY 39 TYSFRLTYVVGIGVALIFGYIQISLWITTAARQTKRIRKQEFHVSVAADIGMFDSCD 98  
 DB 3 TYAV-----YTGIGAGVLIVAVIYQSFMCIAAGRQIHKIRKQFEHAIMROEIGMFDVND 57  
 QY 99 IGEINTRMT-DIDKISDGIKIALLPQNMSTFSGIAGLVKGGKLLVLTSTPLMASAACS 157  
 DB 58 VGEINTRMTDIDKISDGIKIALLPQNMSTFSGIAGLVKGGKLLVLTSTPLMASAACS 117  
 QY 158 SAAACSRVYSLSKELSAVSKAGAAVEEVLSTRTVAFRAQEKELORYONLMDAKDF 217  
 DB 118 SAGIYAKKILTSFEDKELQAAKAGAAVEEVLAVRTVAFGQEKELIRYKHNLEDAKIRIGIRKAI 177  
 QY 218 GKRTTASVSLGAVYFPMNGTYGLAFYGTSLINGEPYTTIGTVLAVFVSIIHSSYICGAVPH 277  
 DB 178 GKRTTASVSLGAVYFPMNGTYGLAFYGTSLINGEPYTTIGTVLAVFVSIIHSSYICGAVPH 235  
 QY 278 GAAPHFETFAIANGAAPHIFQVIDKPKPSIDNEFSTAGKPSIEGTEFKNVSNFVPSRPIKILK 337  
 DB 236 GAAPHFETFAIANGAAPHIFQVIDKPKPSIDNEFSTAGKPSIEGTEFKNVSNFVPSRPIKILK 295  
 QY 338 SIKILGLNLRKSGEYVALVGLNGSGKSTVQVLQRLYDPDGFIVNDENDIRALNRYHRDHIG 397  
 DB 296 DVQILKGLNLRKSGEYVALVGLNGSGKSTVQVLQRLYDPDGFIVNDENDIRALNRYHRDHIG 355  
 QY 398 YRDHIGVSGEPVLFGTITSNKIKYGRDVTDEMERARAEANAYDFIMEPKNKNTLVGEKGAOM 457  
 DB 356 YRDHIGVSGEPVLFGTITSNKIKYGRDVTDEMERARAEANAYDFIMEPKNKNTLVGEKGAOM 415  
 QY 458 EKGAKMSGGOKRIATARALVRNPKILLDEATSLDSEKSAVOALEK 507  
 DB 416 EKGAKMSGGOKRIATARALVRNPKILLDEATSLDSEKSAVOALEK 465

RESULT 9  
 088331 PRELIMINARY: PRT: 1321 AA.  
 AC 088331:  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE P-GLYCOPROTEIN SISTER PROTEIN.  
 GN SPGP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX CHILD S.J., Yeh R.L., Hui D., Ling V.;  
 RA "Taxol resistance mediated by the liver-specific sister gene of p-  
 RT glycoprotein.";

RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS).  
 DR EMBL: AF010597; AAC24753.1; -  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 DR KAP-binding; Transport.  
 KW KAP-binding; Transport.  
 SQ SEQUENCE 1321 AA; 146285 MW; 27F67EC3660080DF CRC64;

Query Match 48.0%; Score 1243.5; DB 11; Length 1321;  
 Best Local Similarity 47.0%; Pred. No. 3.6e-74;  
 Matches 255; Conservative 98; Mismatches 153; Indels 37; Gaps 4;

QY 1 MIIIGILASIVNGACLPMLPLVIGEMSDNLI-----SGCLVQNTMY----- 40  
 DB 62 MIMGGVCALHGMAGPGLIIFGIMDIIFIKYDIERQELTPGKACVANNITVWINSSPHQ 121  
 QY 41 -----SFFRLTYVVGIGVALLIFGYIOISLMTITTAARQTRIRKOPFHS 85  
 DB 122 NMEGTVGGLVDESEMIKFSGIVAGVMVLLIGFQIRLAVITGAGQIRMRKIFRR 181  
 QY 86 VLAODICWFDSCDIGELNTRMT-DIDKISDGIKRIALLFONMSTFGSLAVGLVKGKL 144  
 DB 182 IMREIMCFDCTSVGELNSRPADIIEKINDAIDOLAHFLORMSTAMCGLLIGFRGKL 241  
 QY 145 TLVLTSTPLTMSAASCSRNVISLTSEKELAYSAGAAVEEVLSSIRTVIAFRAQEKEL 204  
 DB 242 TLVLAASPLIGIGANVIGLSIAKFTLELKAIVKAGSIADVLSITVAAFGEKNEV 301  
 QY 205 QRYTONLKDAKDFGIKRTIAKSVLSGAVYFPMNGTYGLAFVGTSLINGEGYITGL 264  
 DB 302 ERYEKNLVFAQRWGIKMGVMVGFYGMWCLIFPCYALAFVYSGTLVLD-EBEYPGTLV 360  
 QY 265 AVFESVHSICIGANVHFETPAIRGAHFIIQVIDIKKSIDNFSTAGYKPESECTV 324  
 DB 361 QIFCLVILAAANNIGHASCSLEIFSTGCSAAATNIFQTIIDROPVTDMSGDGYKLDRIKEI 420  
 QY 325 EFFVSNFVPSRPSIKILKGLNLRKISGEVALVGLNGSGSTVYVOLLORYLDPDGFIM 384  
 DB 421 EFHNVTHYPRPDVKILINDISWIKPEETALVSGSAGSTALQILQIRYIDPCEGMYT 480  
 QY 385 VDENDIRALNVRYRDHIGVYSOEVLFGTTISNNIKYGRDVTDEMERARERANAYDF 444  
 DB 481 LDGHDIRSLNIRWLRDQIGIYEQEVLEFSTTIAENIRFGREDATMEDIVQAKADANAYNF 540  
 QY 445 IMEPNKNFTLVGEKGAOMSGOKORIAARALVRNPKILLIDETSALDSESKAVQAA 504  
 DB 541 IMALPOOPDTLVGEGGSGGOKORAVAIARALRNPKILLIDMATSALEDSEARVOEA 600  
 QY 505 LEK 507  
 DB 601 LNK 603

RESULT 10  
 045721  
 ID 045721 PRELIMINARY: PRT: 1294 AA.  
 AC 045721: 062101;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE C47A10.1 PROTEIN.  
 GN C47A10.1  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Basham V.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten N., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Thierry-Mieg J., Thomas K., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sprout J., Woldman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans."  
 RL Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Basham V.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS).  
 DR EMBL: 293782; CAB07855.1; -  
 DR EMBL: 281484; CAB07855.1; JOINED.  
 DR EMBL: 281484; CAB03973.1; -  
 DR EMBL: 293782; CAB03973.1; JOINED.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR InterPro: IPR00130; Zn\_MiPepdase.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1294 AA; 142490 MW; DD8ED2602FRA484DF CRC64;

Query Match 44.7%; Score 1158; DB 5; Length 1294;  
 Best Local Similarity 45.0%; Pred. No. 1.7e-68;  
 Matches 235; Conservative 95; Mismatches 174; Indels 18; Gaps 4;

QY 1 MIIIGILASIVNGACLPMLPLVIGEMSDNLI-----GCLVQNTMY-----TYSFRL 45  
 DB 46 LAVGIIVSCATGVGLPLMSIINGNVSONFVTLGITFLDPNSTASAKAARAEFSHEVION 105  
 QY 46 TLVYVIGVALLIFGYIOISLMTITTAARQTRIRKQFHSVLAODICWFDSCDIGELNTR 105  
 DB 106 CLKVYVLCGIFPAAGFLASCFMVICERLSNFRQFPFHSVNRQIADWKTSGLTSLNK 165  
 QY 106 MTD-IDKISDGIKRIALLFONMSTFGSLAVGLVKWGLTLVLTSTPLMASAASCSR 164  
 DB 166 LFDNIERYAREGIDKVGGLAFOMMAOPIGFAVAFYDMILTITIMSLSPFMICGLFLAK 225  
 QY 165 NVISLTKELASVSKAGAAVEEVLSSIRTVIAFRAQEKELORYTONLKDAKDFGIKRTIA 224  
 DB 226 LLATAATREAKQYAVAGGIAEVLISIRVIAFNGOEYECRKYEDALEHGKRTGIKKSPL 285  
 QY 225 SKVSLGAVYFPMNGTYGLAFVGTSLINGEGYITGLVLAFFSVIHSYICGAAPVF 284  
 DB 286 IGAGIASFEVLIYASCLAFWGTNFVYSGR--LESGLVLTFFSVMGMSMLGAGGQGF 343  
 QY 285 EFPAIRGAAPFIIOVIRKPSIDNFSTAGYKPESECTGVEEKNSFNPSRPSIKILKG 344  
 DB 344 AITGTRALGAASLVYIDRIPEIDAYSTEGTPSKISGKISVKNKVEFTYPTPADVKILKG 403

[illegible]

OY	1	MIGLISLVAGACLPMPILVIGMS-----DNLSICLVQTMYS---	42
Db	50	LITGVAANVHIGAGPLLAIVLVGGMTVFLRAONSDFVGVGDVNNPGLVPSIDENSE	109
OY	43	-FRLLVYVGIGVAAVLFGYVQISLMTTARQTKRKQFHSVLADQIWGFDSCDICE	101
Db	110	VVKCYIVLVGLVGMFFTSYVQVACFESYARVLVHKLRQNYKALVQLQIOWMEDKQOTG	169
OY	102	LNTWMT-DIDKIISDGIGDKTALLFQNNSTPESIGLAVGLVQWKLTLLVTLSPMLASAA	160
Db	170	LTARLTDLLEVRREGIDGKFFALLVQMFAPLAGYGVGFYSWSMTLVMMGFAPLVLISGA	229
OY	161	ACSRVAVISLTKELSVASKAGAVAEVLSTRITVIAARAQEKELQRTQNLKQAKDGCIR	220
Db	230	KMSKSMARTRVEDEETAVAVAAIAEETFSIRTVHSLNGHRELDREPRYMLAEVGRQIVG	289
OY	221	RTIASVSLGAVVFFPMNGTVLAWMGSGSLINCEPGTIGTVLAVFVSYHSISCGAA	280
Db	290	KYCYGIVGVGSSNLICMSSYALAAWYSGTLIIN-DPFFDRGLITTVFAYLSGTSIGA	348
OY	281	VPHETFAIAGAAFHIEFOVIDKKPSIDNEFSTAGYKDESIEGTVEFKNVSFNYPSPRSIK	340
Db	349	LPHLASFCTAGCAASTVLRVINSHPKIDPYSLEGILVDMNKGDISFQDFVHFRPSRKDIH	408
OY	341	ILKGLNLRKSGEIVLVGLNGSKSTVYVOLLQVLPDPDGFIVNDENDIRALNVKRYRD	400
Db	409	VLKGISLTKAGDKADIKLVAGSSGCKSTVNLLOFRPYDTKGRVLIDVDLRENVVLSRE	468
OY	401	HIGVVSQEPVLEFGTTISNNIKYKGRDDVDYDEMEARAAREANAYDPIMEFPKKEFVLVSEKG	460
Db	469	QIGIVSQEPVLEFGDTIYNINIKMGNEHATHQVYVACKMANANDEIRKLDPGCTGRVSEKG	528
OY	461	AQMSGQKQRIARALVARNPKLILDEATSSALDSSEKSAVQAALKEK	507
Db	529	VQLSGQKQRIARALVARNPKLILDEATSSALDTEAREVQALDQ	575
RESULT	12		
O9V626			
O9V626		PRELIMINARY;	PRT; 1313 AA.
AC	O9V626;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	MDR50 PROTEIN.		
GN	MDR50 OR CG8523.		
OS	Drosophila melanogaster (Fruit fly)		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;		
OC	Ephydrioidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RP	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RC	MEDLINE=20196006; PubMed=107311132;		
RC	Adams M.D., Celisliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RC	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RC	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RC	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RC	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RC	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RC	Abdl J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,		
RC	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RC	Beeoon K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,		
RC	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,		
RC	Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RC	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RC	de Pablos B., Delcher A., Deng Z., Davis A.D., Dew I., Dietz S.M.,		
RC	Dosdon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.		
RC	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RC	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RC	Glodet A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,		
RC	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Modyaty G., Morris J., Mostrefti A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitkas R., Tector G., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS).  
 CC EMBL: AE003815; AAF58271.1; -;  
 DR FlyBase; Fggn0010241; Mdr50.  
 DR InterPro; IPR003593; AAA.  
 DR InterPro; IPR001140; ABC\_transporter\_tmam.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR001687; ATP\_GTP\_A.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1313 AA; 145277 MW; 3BB9DEA690AADB58 CRC64;

Query Match 41.1%; Score 1065; DB 5; Length 1313;  
 Best Local Similarity 42.8%; Pred. No. 2.8e-62;  
 Matches 227; Conservative 97; Mismatches 178; Indels 28; Gaps 7;

QY 2 ILGIASLVNGACLPMPPLVIGEMSDNLI--SGCLVQNTNY-----SFFRLTL 47  
 DB 89 VIGLSLVNATGTLTPANSLIFGNLANMDLGLGLESKSTRADASTLLDKVRPFL 148  
 QY 48 YVVGIGVAALIFGYIQISLMTITTAARQTKRIKQPFHSHVLAQDGMFDSICDELTRMT 107  
 DB 149 QNVIYIGIMLVCSYSLTTCNVAAHSQILITRSKFFRSILHODKMWDFNDSGEVASRMN 208  
 QY 108 -DIDKISDGIQDKIALLFQNNKSTFSGILAVGLVGMKLTLYTSTPLIMAS----AAAC 162  
 DB 209 EDLSKMDGLAEKVMFVHVLVAFGSLVLAFAVGMQLSLVCLTSLPLTFAMGLVAVAT 268  
 QY 163 SRWVISTSKELSAVSKAGVAEEVLSIRIVYAFRAQEKELQRYTONLKAOKFGIKRT 222  
 DB 269 SR----LAKREVIMYAGAAVAAVEGALSIGIRVKAEGEAEVAAYKRRVAAKILNTKRN 324  
 QY 223 IASKVSLGAVFFMNGTYGLAFWGTSLILNG--EP---GYTIGTVAFFSVHSYCI 277  
 DB 325 MFSIGIGLIMFTIYASTALAFWGVGLVIGKIHPRAYENTDAGTMTITVFFSVAMGSMTI 384  
 QY 278 GAAPVHEFTALANGAFHIFQVLDKRPISDNFSTAGYKPEISIEGTVEFKVNSFNPSP 337  
 DB 385 GMAPVTEARGIAKACAKVFIIEQIPEINPIDEGKKLMEPLTTEFKVEFVQYPTRP 444  
 QY 338 SIKLKLRKIKSGEVALVGLNGSKSTVYOLLQRLYPPDDGFIMVDENDIALNVRH 397  
 DB 445 EVSLTINKLTKIRHGQVALVGLPGSGCKSTCIOLVORRYDQAGNLNGLNGLDLDL 504  
 QY 398 YRHQIGVSOEPLVFTTISNNIKYGRDDVTDDEMERARARANAYDFIMEPNKFN 457  
 DB 505 LRSIGVVGQEPFLFATSIYENIRKREDATREIEAANAANAIAIFPKLPKGYDILVG 564  
 QY 458 EKAQMSGGOKORIAIARALVRNPKIILDEATSAIDSEKSAVOALEK 507

DB 565 EKAQMSGGOKORIAIARALVRNPKIILDEATSAIDSEKSAVOALEK 614  
 RESULT 13  
 ID 061301 PRELIMINARY; PRT; 1275 AA.  
 AC 061301;  
 DT 01-AUG-1998 (Tremblrel. 07, Created)  
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE P-GLYCOPROTEIN.  
 GN Pep-A.  
 OS Haemonchus contortus.  
 OC Eukaryote; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Trichostrongylidae; Haemonchidae; Haemonchinae; Haemonchus.  
 OX NCBI\_TaxID=6289;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9826176; PubMed=9566525;  
 RA Xu M., Molento M., Blackhall W., Ribeiro P., Beech R., Pritchard R.;  
 RT "Ivermectin resistance in nematodes may be caused by alteration of p-  
 glycoprotein homolog.";  
 RL Mol. Biochem. Parasitol. 91:327-335(1998).  
 CC -1- SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC  
 TRANSPORTERS).  
 CC EMBL: AF003908; AAC38987.1; -;  
 DR HSSP; P13569; INBD.  
 DR InterPro; IPR003593; AAA.  
 DR InterPro; IPR001140; ABC\_transporter\_tmam.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR001687; ATP\_GTP\_A.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 2.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1275 AA; 142203 MW; 7A39D3836B75787C CRC64;

Query Match 40.5%; Score 1050.5; DB 5; Length 1275;  
 Best Local Similarity 42.6%; Pred. No. 2.5e-61;  
 Matches 223; Conservative 94; Mismatches 186; Indels 21; Gaps 5;

QY 4 GILASLVNGACLPMPPLVIGEM-----SDNLSGCL-----VQNTYSFF-----R 44  
 DB 54 GLAPCVHAGAFSVLGYIVLGMFTVFLRAQSEFVLGVSHPDGLPALTFKEEDTLVRR 113  
 QY 45 LTLVYVIGVAALIFGYIQISLMTITTAARQTKRIKQPFHSHVLAQDGMFDSICDELNT 104  
 DB 114 YCLVYLGIGFMFATSYQIYICWETFERITHKRLKITYLKAIRQOISWFDIQGTGLTA 173  
 QY 105 RMT-DIDKISDGIQDKIALLFQNNKSTFSGILAVGLVGMKLTLYTSTPLIMASAAACS 163  
 DB 174 RLTDLLEVRGLDQKLSLFTQMSAFVAGFCVGFAYMSMTLVMVAVAPVITSAAMMS 233  
 QY 164 RMVISTSKELSAVSKAGVAEEVLSIRIVYAFRAQEKELQRYTONLKAOKFGIKRTI 223  
 DB 234 KIVATRVQDEETAVAGALAEIEFTSSIRVHSGIKHRELTFREALEKROGLVYKF 293  
 QY 224 ASKVSIGAVFFMNGTYGLAFWGTSLILNGEPGTTIGTVLAFVFSYHSYICGAAPVH 283  
 DB 294 YMGVGVGQGCQIVSYVALAFWGTSLIIN-DPALDRGRIFTVFPAVMSGSAALGTCLPH 352  
 QY 284 FETFAIANGAFHIFQVLDKRPISDNFSTAGYKPEISIEGTVEFKVNSFNPSPSIKILK 343  
 DB 353 LNTISIAAGAVRSVLSVNSRPKIDPVSLDGIVLNMMGSLTFKRVHFSYFSRRLQILK 412  
 QY 344 GLNLRKSGEVALVGLNGSKSTVYOLLQRLYPPDDGFIMVDENDIALNVRH 403  
 DB 413 GVSILQVSAQKIALVGLSSGCKSTVNTVLLRFYDPTBRKVITIDIDVCDLWVQKLRQIG 472  
 QY 404 VVSOEPLVFTTISNNIKYGRDDVTDDEMERARARANAYDFIMEPNKFN 463

Dd		:     :     :         :		
Oy	464 SGGOKORIATARAALRNPKILLIDEATSALDSSEKSAVOAALEK	507		
Dd	533 SGGOKORAIARAIKINPRILLIDBATSALDTEASTIYOEALEK	576		
	RESULT 14			
	09LGX1 PRELIMINARY; PRT: 1285 AA.			
AC	09LGX1:			
Dt	01-OCT-2000 (TREMBLrel. 15, Created)			
Dt	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
Dt	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
De	SIMILAR TO ARABIDOPSIS THALIANA CHROMOSOME 2.			
Oe	Oryza sativa (Rice).			
Oc	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Oc	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Ox	Einhartoidae; Oryzaceae; Oryza.			
RN	NCB1_TaxID=4530;			
RP	[1]			
RC	SEQUENCE FROM N.A.			
RA	STRAIN=cv. NIPPONBARE;			
RT	Sasaki T., Matsumoto T., Yamamoto K.;			
RL	"Oryza sativa japonbare(Ga3) genomic DNA, chromosome 1, PAC			
CL	clone:p0706805."			
CC	Submitted (JUN-2000) to the EMBL/Genbank/DDBJ databases.			
DR	-1 SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC			
DR	TRANSPORTERS).			
DR	EMBL, AP002482; BAA96612.1; -.			
DR	InterPro: IPR0033593; AAA.			
DR	InterPro: IPR001140; ABC_transporter_tmam.			
DR	InterPro: IPR003439; ABC_transporter.			
DR	InterPro: IPR001687; ATP_GTP_A.			
DR	Pfam: PF00664; ABC_membrane; 2.			
DR	Pfam: PF00005; ABC_tran; 2.			
DR	SMART: SM00382; AAA; 2.			
DR	PROSITE: PS00211; ABC_TRANSPORTER; 2.			
KW	ATP-binding; Transport.			
SO	SEQUENCE 1285 AA; 138049 MW; 2D45FEC7P956A1A0 CRC64;			
	Query Match 39.9%; Score 1033; DB 10; Length 1285;			
	Best Local Similarity 40.9%; Pred. No.3.6e-60;			
	Matches 208; Conservative 112; Mismatches 185; Indels 4; Gaps			
OY	1 MILGLIASLVNGACLPLMLPVIGENSDNLISGLVQNTNTYSFFRTLTVYGIVGAALIFG	60		
Dd	71 MALGTGLGVANCAALPEFTVLFGNLIDAFCAGMGHVDVNVRMSVSLRFYLAIASAVS	130		
OY	61 YIOISLIITTAARKRKRIKOPFSVLADIGMDEPS-CDIGELNRM-T-DIKIDSIGICD	118		
Dd	131 FVOYTGMWTGERQAARIKNLTLYKLROIAFEFKYTYTGSGVGMSDVLLIDAMEE	190		
OY	119 KIALLFOUMSFESISLAGLVKGWLTLVTSTSPDLIMASAACSRLVSIITSKELSAYS	178		
Dd	191 KYGRKIQLVLPFLGFIFYAFNQGWILLTVMATTIPVIYAGVAHSVNVAKMASLDQAATA	250		
OY	179 KAGAFAVEEVLSIRFVIAFRAOEKELORYTONLKDAKDGIKRTTASKEYSLGAYVFENG	238		
Dd	251 ESSVVVEOTIGISRIVASFTEGKAQVEKYNSLKSAYSGSVREGILAAGIMGTMVLLFC	310		
OY	239 TYGLAPFWGTSLNLGEFGYTIGTVLAVFVSFIHSXYCSLGAAVPHPFETPAIRGAATHF	298		
Dd	311 GSTSIDIWGAKLILT-LKGYTGAKAMNYIFEAVLTSSSLAGQASPWKAPAAGQAAAAYKF	368		
OY	299 QVIDKKPSIDNSTAGRYKPESIEGTVEEFKNVSNFPSPRSIKTLGLNLRIRKSETVALY	358		
Dd	369 ETINKKPELDASTYGMKPPDDIRGOIERPDVYFESTPTRRDEDIFRGFSISTPSGTTVALY	428		
OY	359 GLNGSKSTVVOQLRLVDPDDGFMVDENDIRALNVRHRDHIGVSGEPVLFEGTTISN	418		

Dd	429	GQSSGKSTVSIILERYDDQLDGLVDLIDGNLNKEPDLRNRISKTIGLYSEPLVFAASIKIE	486
Oy	419	NKKGRDVDTDEEMARAREANAYDFIMEPNKNFTLVGEKGAGMSGGOKRIATARALY	478
		:    :            :      :       :	
Dd	489	NNAVGKNATQDEIRAAAEALNLANSKFIDKMPOGLDITVSVEHGTQLSGGOKRIATARAIL	548
Oy	479	RNPKITLIDEATSALDSSEKSVAOALEK 507	
		:::  :       ::  :      :	
Dd	549	KDPRIILLDEATSLDAESERIVOEALDR 577	
RESULT	15		
ID	Q9V616	PRELIMINARY; PRT; 1279 AA.	
AC	Q9V616;		
DT	01-MAY-2000 (TREMBlrel_13, Created)		
DT	01-MAY-2000 (TREMBlrel_13, Last sequence update)		
DT	01-JUN-2001 (TREMBlrel_17, Last annotation update)		
DE	MOR49, PROTEIN.		
GN	MOR49 OR CG3879.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
CC	Ephyridioidea; Drosophilidae; Drosophila.		
OX	NCBI_Taxid=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY.		
FX	MEDLINE=20196006; Pubmed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Burton R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abdil J.F., Agbayani A., An H.-C., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeoon K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Casley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doug L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garz N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hosain D., Houston K.A., Howland T.J., Wei M.-H., Iibegwan C.,		
RA	Jaitai M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Kethum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., Mcleod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Murzyn D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusseren D.R., Pacleb J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spiet E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Welnstock G.M., Weissbach J.,		
RA	Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.).		
RT	"The genome sequence of Drosophila melanogaster.";		
RL	Science 287:2185-2195(2000).		
CC	-I SIMILARITY: TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC		
CC	TRANSPORTERS).		
DR	EMBL; AE003820; AAF58437.1; -		
DR	FLYBase; FBgn0004512; Md419.		
DR	InterPro; IPRO003593; AAA.		
DR	InterPro; IPRO001140; ABC_transporter_tmem.		



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:34:58 ; Search time 31.47 Seconds  
(without alignments)  
367.547 Million cell updates/sec

Title: US-09-873-409-8

Perfect score: 2592  
Sequence: 1 MILGILASLVNGACLPMLPL.....SESKSAVQALEKDPFRYSF 514

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1576	60.8	1275	3 US-09-120-513-2	Sequence 2, Appli
2	1576	60.8	1275	4 US-09-450-105-2	Sequence 2, Appli
3	1576	60.8	1280	2 US-08-752-447-2	Sequence 2, Appli
4	1568	60.5	1280	2 US-08-583-276-19	Sequence 10, Appli
5	1568	60.5	1280	6 5206352-4	Patent No. 5206352
6	1553.5	59.9	1279	2 US-08-784-649A-2	Sequence 2, Appli
7	975.5	37.6	1408	1 US-08-612-521-2	Sequence 2, Appli
8	959	37.0	1349	2 US-08-612-734B-2	Sequence 2, Appli
9	937	36.1	1334	4 US-08-996-545-2	Sequence 2, Appli
10	937	36.1	1334	4 US-09-328-320-2	Sequence 2, Appli
11	920	35.5	1307	1 US-08-385-246C-2	Sequence 2, Appli
12	762.5	29.4	1308	2 US-08-996-644-2	Sequence 2, Appli
13	762.5	29.4	1308	3 US-09-352-552-2	Sequence 2, Appli
14	669.5	25.8	791	1 US-08-394-880B-2	Sequence 2, Appli
15	639.5	24.7	1302	1 US-08-232-537-2	Sequence 2, Appli
16	635	24.5	686	4 US-09-061-764A-15	Sequence 15, Appli
17	629	24.3	653	4 US-09-061-764A-2	Sequence 2, Appli
18	609	23.5	327	1 US-08-463-092B-9	Sequence 9, Appli
19	609	23.5	327	2 US-08-460-907B-9	Sequence 9, Appli
20	598	23.1	748	4 US-09-061-764A-19	Sequence 19, Appli
21	515.5	19.9	707	3 US-08-772-270A-4	Sequence 4, Appli
22	499.5	19.3	694	3 US-08-895-522-4	Sequence 4, Appli
23	499.5	19.3	694	3 US-09-195-391-4	Sequence 4, Appli
24	497.5	19.2	711	3 US-08-772-270A-12	Sequence 12, Appli
25	472.5	18.2	747	3 US-08-895-522-1	Sequence 1, Appli
26	472.5	18.2	747	3 US-09-195-391-1	Sequence 1, Appli
27	463	17.9	694	2 US-08-895-522-3	Sequence 3, Appli

28	463	17.9	694	3 US-09-195-391-3	Sequence 3, Appli
29	397	15.3	110	1 US-08-466-886-22	Sequence 22, Appli
30	397	15.3	110	4 US-08-466-817-22	Sequence 22, Appli
31	392	15.1	110	1 US-08-466-886-20	Sequence 20, Appli
32	392	15.1	110	4 US-08-466-817-20	Sequence 20, Appli
33	381.5	14.7	109	1 US-08-466-886-24	Sequence 24, Appli
34	381.5	14.7	109	4 US-08-466-817-24	Sequence 24, Appli
35	354	13.7	1477	3 US-08-492-459-10	Sequence 10, Appli
36	354	13.7	1477	3 US-08-423-752-10	Sequence 10, Appli
37	354	13.7	1477	3 US-08-945-994-3	Sequence 3, Appli
38	354	13.7	1477	4 US-08-716-873-24	Sequence 24, Appli
39	354	13.7	1477	4 US-09-368-431-24	Sequence 24, Appli
40	353.5	13.6	400	4 US-08-961-083-190	Sequence 190, Appli
41	350.5	13.5	1437	3 US-09-061-400-2	Sequence 2, Appli
42	350.5	13.5	1437	2 US-09-001-273-2	Sequence 2, Appli
43	350.5	13.5	1453	4 US-08-843-459A-2	Sequence 2, Appli
44	346.5	13.4	1621	4 US-08-972-927-3	Sequence 3, Appli
45	345	13.3	1528	1 US-08-463-092B-6	Sequence 6, Appli

#### ALIGNMENTS

RESULT 1  
US-09-120-513-2  
Sequence 2, Application US/09120513  
Patent No. 6025160  
GENERAL INFORMATION:  
APPLICANT: Brun, Kimberly  
APPLICANT: Chenery, Richard  
APPLICANT: Ellens, Harma  
APPLICANT: Field, John  
APPLICANT: Yue, Lin  
TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
TITLE OF INVENTION: SEQUENCES ENCODING RAT MPR182 AND  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY:  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/120.513  
FILING DATE: 22-JUL-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: King, William T  
REGISTRATION NUMBER: 30,954  
REFERENCE/DOCKET NUMBER: GP50008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5015  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1275 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-120-513-2  
Query Match 60.8% Score 1576: DB 3: Length 1275;  
Best Local Similarity 58.4% Pred. No. 2e-146;



```

Matches 308; Conservative 92; Mismatches 105; Indels 22; Gaps 4;

OY 1 MTLGLASLVNGACPLPMLPYLVGKSDNL-----SGCLVQNTNTYSFRL--- 45
    |||:::|||||::|::|::|::|
Db 49 MALGTALAIIGTLLPLMLVFGVMTDSTFOAETRLPVSVMQSSINQTVSDSLEED 108
OY 46 ----TLYVVGIGVVALIFGYIOISLWTTAARQTRIRKOPFHSVLADIGFSDCDIGE 101
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 109 MAMTAYTYTGTGAGVLYIAYTVQVSLMCLAGQRIHKIRQKFFHALIMQDIEGFVDNDAGE 168
OY 102 LNTFRMT-DIDRISDIDGKIALLFQNMSTFSIGLAVGVKGLTVLSTPLIMASA 160
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 169 LNTRLTDDVSKINISGIDKLGMPQSITTFEAGFIIGFISGKLLTVLTVLAVSPLIGSSA 228
OY 161 ACSRMVJSLTSKELSAYSKAGAVAEVUSITRTVYAFRAQKEIORYONLKDADFQIK 220
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 229 MMAKLVLTFTTKKELOAVYAKAGVAEEVLAIRTVAFGQKKEIERYNKNLEAKRVGIK 288
OY 221 RTIASKVSGLAVYFPMNCTYGLAFMYGTSILNIGEPGYTIGTVLAVFVSYHSSVCIGA 280
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 289 KAITANISIGLAVLLVYASVALAFMYGTSVLSYLSNE--YSIGVLTFFYSILGTESIGHL 346
OY 281 VPHEFTFIARGAAEHIPOVIDKPSIDINEFSPAGYKPEISCTEVEKNVSFNVPSPSIK 340
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 347 APNTEAFANAGCAVAEIRKIIDNESIDSFSKGRHKPDSIMKNLEKXNYFNVPSPSEVK 406
OY 341 ILKGLNLIRKSGEVALVGLNCSGKSTVYQLLRLYDPDDGFIWDENDIRALNVRHYVD 400
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 407 ILKGLNLEVYKSGQFVALVGNCSGCKSTVQLLRLYDPDIEGEVSIIDQODIRFINVRYLKE 466
OY 401 HIGVVSQEPVLEGTITSNNIKKYGRDQVDDEMERARARENADVIMEPNKENTLVGEGK 460
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 467 IIGVVSQEPVLEFATTIAEKIRIGRENVTMDELEKAVKEANADVIMKLDPKHNFTLVGEGK 526
OY 461 AQMSGGQKQRIAIARALYRNPKILLIDETSALDSESSKSAVOAALEK 507
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 527 AQLSGGQKQRIAIARALYRNPKILLIDETSALDSESSAIVOAALEK 573

RESULT 2
US-09-450-105-2
; Sequence 2, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Cheney
; APPLICANT: Harna Ellens
; APPLICANT: John Anthony Field
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDRI1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1275
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-450-105-2

```

```

Query Match Similarity      60.8%  Score 1576: DB 4:  Length 1275;
Best Local Similarity      58.4%  Pred No. 26-116;
Matches 308; Conservative  92; Mismatches 105; Indels 22; Gaps 4

QY 1 M I G T I A S L V N G A C P L M V Y L G E N S D N I L ----- S G C L V Q N T Y S F F L --- 45
DB 49 M A L G T I A A I I H T L P L L V A F G Y M D S T F Q A E R T L P S V T N O S E I N S Q T V D S S L E E D 108

```

[illegible]

```

1      RESULT      3
2      US-08-752-447-2
3      ; Sequence 2, Application US/08752447
4      ; Patent No. 5994088
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Mechtner, Eugene
7      ; APPLICANT: Roninson, Igor B
8      ; TITLE OF INVENTION: Methods and Reagents for Preparing and
9      ; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
10     ; NUMBER OF SEQUENCES: 2
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESS: McDonnell Boehnen Hulbert & Berghoff Ltd.
13     ; STREET: 300 South Wacker Drive, Seventh Floor
14     ; CITY: Chicago
15     ; STATE: Illinois
16     ; COUNTRY: USA
17     ; ZIP: 60606
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Floppy disk
20     ; COMPUTER: IBM PC compatible
21     ; OPERATING SYSTEM: PC-DOS/MS-DOS
22     ; SOFTWARE: Patent Release #1.0, Version #1.30
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/08/752.447
25     ; FILING DATE: 15-NOV-1996
26     ; CLASSIFICATION: 435
27     ; ATTORNEY/AGENT INFORMATION:
28     ; NAME: No. 5994088nan, Kevin E
29     ; REGISTRATION NUMBER: 35,303
30     ; REFERENCE/DOCKET NUMBER: 95,1121
31     ; TELECOMMUNICATION INFORMATION:
32     ; TELEPHONE: 312-913-0001
33     ; TELEFAX: 312-913-9808
34     ; INFORMATION FOR SEQ ID NO: 2:
35     ; SEQUENCE CHARACTERISTICS:
36     ; LENGTH: 1280 amino acids
37     ; TYPE: amino acid
38     ; TOPOLOGY: linear
39     ; MOLECULE TYPE: protein
40     ;
41     ; US-08-752-447-2

```

```

Query Match          60.8%; Score 1576; DB 2; Length 1280;
Best Local Similarity 57.9%; Pred. No. 2e-146;
Matches 305; Conservative 101; Mismatches 99; Indels 22; Gaps 5;

Oy 1 MILGTLASLVNGACPLPLMPVLVIGMSD-----NLISGCLVOT--NTYSFF----- 43
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 MVVGTLLAIIHGAGLPLMLVFGEMTDIFPANGNLEDLMSNTNRSNDINDGCFNNLEED 110
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 44 --RLTVYVGIGVAAIFGYIQISLMTITTAARQTRIRKQEFHVSVAODIGMFCSDIGE 101
   | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 111 MTRAYVYSGIAGVLAIVQVSWFCCLAAGQRIHKIRKQEFHAIHQEIGMFDVHDVGE 170
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 102 LNTWMT-DIDKISDGIKIALLFQNMSTFSGIAGLVKQKLTLYLSTSPILMASAA 160
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 LNTRLTDVSKINEYIGDKIGMFOSMATFFTGIVGTRGKLTLYLTAISPVLGISA 230
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 161 ACSRWVILTSKELSAVSKAGAAVEEVSIRTVIAFRAQEKELORYONKDKADDFGIR 220
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 VMAKILSFTEKELLATAKAGAAVEEVLAIKRTVIAFGQKKELEKYNKLEAKRIGIK 290
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 221 RTIASVSLGAVYFPMNGTYGLAFWGTSLINBPGYTGTVLAVFVSIIHSSYIGAA 280
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 291 KAITANISIGAFLLIVASYAVAFWYGTTLVSGE--YSIGVLLVFPSVLIGAFVGOA 348
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 281 VPHEFTAIARGAAPHIPOVIDKRPISDNFSTAGYKPSISGTYVEFKVNSFNYSRPSIK 340
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 SPSEAFANAGAAVEIFKIIDNKRPSIDYSKSGHKPNKINGLEFRVNHFSYPSRKEVK 408
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 341 ILKGLNLRKSGEYVALVGLNGSGSTVVOQLRLYDDDFIMVDENDIRALNRYHRD 400
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 409 ILKGLNLRKSGEYVALVGLNGSGSTVVOQLRLYDDDFIMVDENDIRALNRYHRD 468
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 401 HIGVSOEPLVFGTTISNNIKYGRDVTDEEMERARANAYDFIMEPNKFNLTLYGEGK 460
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 469 IIGVSOEPLVFGTTIAENIRGRONVTMDEIEKAVKANAYDFIMKLPKHFDTLYGEBG 528
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 461 AOMSGGOKORIAIARALVRNPKIILDBATSALDESSESAVOALEK 507
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 529 AOMSGGOKORIAIARALVRNPKIILDBATSALDESSESAVOVALDK 575
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-08-583-276-19
; Sequence 19, Application US/08583276
; Patent No. 583/536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Giffillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DNA V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: 08/332,444
FILING DATE: 31-OCT-1994
APPLICATION NUMBER: 07/887,712
FILING DATE: 22-MAY-1992
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1280 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
US-08-583-276-19

Query Match          60.5%; Score 1568; DB 2; Length 1280;
Best Local Similarity 57.9%; Pred. No. 1.3e-145;
Matches 305; Conservative 99; Mismatches 101; Indels 22; Gaps 5;

Oy 1 MILGTLASLVNGACPLPLMPVLVIGMSD-----NLISGCLVOT--NTYSFF----- 43
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 51 MVVGTLLAIIHGAGLPLMLVFGEMTDIFPANGNLEDLMSNTNRSNDINDGCFNNLEED 110
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 44 --RLTVYVGIGVAAIFGYIQISLMTITTAARQTRIRKQEFHVSVAODIGMFCSDIGE 101
   | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 111 MTRAYVYSGIAGVLAIVQVSWFCCLAAGQRIHKIRKQEFHAIHQEIGMFDVHDVGE 170
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 102 LNTWMT-DIDKISDGIKIALLFQNMSTFSGIAGLVKQKLTLYLSTSPILMASAA 160
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 LNTRLTDVSKINEYIGDKIGMFOSMATFFTGIVGTRGKLTLYLTAISPVLGISA 230
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 161 ACSRWVILTSKELSAVSKAGAAVEEVSIRTVIAFRAQEKELORYONKDKADDFGIR 220
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 231 VMAKILSFTEKELLATAKAGAAVEEVLAIKRTVIAFGQKKELEKYNKLEAKRIGIK 290
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 221 RTIASVSLGAVYFPMNGTYGLAFWGTSLINBPGYTGTVLAVFVSIIHSSYIGAA 280
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 291 KAITANISIGAFLLIVASYAVAFWYGTTLVSGE--YSIGVLLVFPSVLIGAFVGOA 348
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 281 VPHEFTAIARGAAPHIPOVIDKRPISDNFSTAGYKPSISGTYVEFKVNSFNYSRPSIK 340
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 SPSEAFANAGAAVEIFKIIDNKRPSIDYSKSGHKPNKINGLEFRVNHFSYPSRKEVK 408
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 341 ILKGLNLRKSGEYVALVGLNGSGSTVVOQLRLYDDDFIMVDENDIRALNRYHRD 400
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 409 ILKGLNLRKSGEYVALVGLNGSGSTVVOQLRLYDDDFIMVDENDIRALNRYHRD 468
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 401 HIGVSOEPLVFGTTISNNIKYGRDVTDEEMERARANAYDFIMEPNKFNLTLYGEGK 460
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 469 IIGVSOEPLVFGTTIAENIRGRONVTMDEIEKAVKANAYDFIMKLPKHFDTLYGEBG 528
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 461 AOMSGGOKORIAIARALVRNPKIILDBATSALDESSESAVOALEK 507
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 529 AOMSGGOKORIAIARALVRNPKIILDBATSALDESSESAVOVALDK 575
   |||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
5206352-4
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO: 4:

```

LENGTH: 1280  
5206352-4

Query Match 60.5%; Score 1568; DB 6; Length 1280;  
Best Local Similarity 57.9%; Pred. No. 1.3e-145;  
Matches 305; Conservative 99; Mismatches 101; Indels 22; Gaps 5;

1 MILGILASLVNGACPLMPVLVGEKSD-----NLISGCLVOT--NTYSFE----- 43  
DB 51 MVTGLAIIHAGGLPLMLVFGEMDIFANAGNEDLMSNTNRSDINDTGFENLDE 110  
44 --RLTLVYVIGVVAALIFGYIOISLMTITPAAROTKRIKOFHVSVAADIGFDSCDIGE 101  
DB 111 MTRVAVYSGIGAVLVAAYIOVSFMCCLAAGROIHKIRKOFPAIRKOEIGWFDVHDYGE 170  
102 LNTBMT-DIDKISDGIIDGIALLFQNMSTFSGLAVALGKWLTLVTLSTPLIMASAA 160  
DB 171 LNTRLTDVSKINQVIGDKIGMFQSMATFETGIVGTGKWLTLVLAISFVLGLSAA 230  
161 ACSRWVISTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELQRYTONLKDAPFGIK 220  
DB 231 VAKIISSTOKELLAYARAGAAVEVLAIKRVIAFGGKKELERYNNKLEAKRIGIK 290  
221 RTIASKVSLGAVYFPMNGTGLAFWYGTSLILNGEPYTGIVLAVFVSYSYICGAA 280  
DB 291 KAITANISIGAEFLIYASALAFWYGTTLVLGSE--YSIGQVLTVFVSYLIGAFSVGQA 348  
281 VPHFEFATARGAAPHIFQYIDKRPISIDNSTAGYKPESTEGTVERKNVSFNPSPRSIK 340  
DB 349 SPSEAFANARGAAVEFLIKIDNKPISIDSKSGHKPDNKGMLERFNHVSFSPSRKRVK 408  
341 ILKGLMLRISGEVVALVGLNGSGKSTVVOQLQRLYDPDDGFIIMVNDIRALNVRHYRD 400  
DB 409 ILKGLMLKVOGQTVLAVNGSGKSTVQMLQRLDPTGEMVSVQGDQIRITINRFLRE 468  
401 HIGVSOEPLVFGTTISNNIKYGRDVTDEMERARAEANAYDFIMEFPKFTLVGEGK 460  
DB 469 IIGVSOEPLVFAITTAENIRYGRVNTMDEIEKAAVEANAYDFIMKLPKRPFTLVGEGK 528  
461 AOMSGGOKRIATARALVRNPKLILDEATSAIDSEKSAVOALEK 507  
DB 529 AOLSGGOKRIATARALVRNPKLILDEATSAIDSEKSAVOALEK 575

RESULT 6  
US-08-784-649A-2  
Sequence 2, Application US/08784649A  
Patent No. 5830697  
GENERAL INFORMATION:  
APPLICANT: SIKIC, Branlimir I  
APPLICANT: Chen, Gang  
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,649A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J

REGISTRATION NUMBER: Reg. No. 5830697 36,677  
REFERENCE/DOCKET NUMBER: 06037/007001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-322-5070  
TELEFAX: 415-854-0875

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1279 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-784-649A-2

Query Match 59.9%; Score 1553.5; DB 2; Length 1279;  
Best Local Similarity 57.7%; Pred. No. 3.4e-144;  
Matches 304; Conservative 100; Mismatches 100; Indels 23; Gaps 6;

1 MILGILASLVNGACPLMPVLVGEKSD-----NLISGCLVOT--NTYSFE----- 43  
DB 51 MVTGLAIIHAGGLPLMLVFGEMDIFANAGNEDLMSNTNRSDINDTGFENLDE 110  
44 --RLTLVYVIGVVAALIFGYIOISLMTITPAAROTKRIKOFHVSVAADIGFDSCDIGE 101  
DB 111 MTRVAVYSGIGAVLVAAYIOVSFMCCLAAGROIHKIRKOFPAIRKOEIGWFDVHDYGE 170  
102 LNTBMT-DIDKISDGIIDGIALLFQNMSTFSGLAVALGKWLTLVTLSTPLIMASAA 160  
DB 171 LNTRLTDVSKINQVIGDKIGMFQSMATFETGIVGTGKWLTLVLAISFVLGLSAA 230  
161 ACSRWVISTSKELSAVSKAGAAVEVLSSIRTVIAFRAOEKELQRYTONLKDAPFGIK 220  
DB 231 VAKIISSTOKELLAYARAGAAVEVLAIKRVIAFGGKKELERYNNKLEAKRIGIK 290  
221 RTIASKVSLGAVYFPMNGTGLAFWYGTSLILNGEPYTGIVLAVFVSYSYICGAA 280  
DB 291 KAITANISIGAEFLIYASALAFWYGTTLVLGSE--YSIGQVLTVFVSYLIGAFSVGQA 348  
281 VPHFEFATARGAAPHIFQYIDKRPISIDNSTAGYKPESTEGTVERKNVSFNPSPRSIK 340  
DB 349 SPSEAFANARGAAVEFLIKIDNKPISIDSKSGHKPDNKGMLERFNHVSFSPSRKRVK 407  
341 ILKGLMLRISGEVVALVGLNGSGKSTVVOQLQRLYDPDDGFIIMVNDIRALNVRHYRD 400  
DB 409 ILKGLMLKVOGQTVLAVNGSGKSTVQMLQRLDPTGEMVSVQGDQIRITINRFLRE 467  
401 HIGVSOEPLVFGTTISNNIKYGRDVTDEMERARAEANAYDFIMEFPKFTLVGEGK 460  
DB 469 IIGVSOEPLVFAITTAENIRYGRVNTMDEIEKAAVEANAYDFIMKLPKRPFTLVGEGK 527  
461 AOMSGGOKRIATARALVRNPKLILDEATSAIDSEKSAVOALEK 507  
DB 528 AOLSGGOKRIATARALVRNPKLILDEATSAIDSEKSAVOALEK 574

RESULT 7  
US-08-612-521-2  
Sequence 2, Application US/08612521  
Patent No. 5786463  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B  
APPLICANT: Skatrud, Paul L  
APPLICANT: Thornewell, Susan J  
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patent Division/AEH  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: USA



QY 466 GOKORIAIARALVNRPKILLDEATSAIDSEKSAVOALEK 507  
Db 588 GOKORIAIARIVSDPKILLDEATSAIDTRKSEGVQALDK 629

## RESULT 9

US-08-996-545-2  
; Sequence 2, Application US/08996545  
; Patent No. 5928898  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,545  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1334 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-996-545-2

Query Match 36.1%; Score 937; DB 2; Length 1334;

Best Local Similarity 41.2%; Pred. No. 2.7e-83;

Matches 194; Conservative 96; Mismatches 169; Indels 12; Gaps 3;

QY 47 LYYVGIGVAAALIFGYIOISLITTAAROTKRIKQFHSVLAODIGMFDSCDDELNTRM 106  
Db 153 LYYVLLIGGEVTVYVSTGVGTYGEHATOKIRREYLESILRONIGYFDKIGAGEVTRI 212  
QY 107 T-DIDKISDGIKIALLFQNMSTFSGIAGLVKWKLTFTVLTSTPLIMASAAACSRM 165  
Db 213 TADTNLIQDGISEKVGTLTALATFVTAFTIAYYKMKLALICSTIVALVLTNGGSSQF 272  
QY 166 VISLTSKELSAVSKAGAVAEVLSSIRTVIAFRAOEKELQRYTONLKDADFGIKRTIAS 225  
Db 273 IIRKSKSLDSYGAGVYAEVVISIRNATFQTDOKIAKQYEVHDEAEKMGTRKNOIYM 332  
QY 226 KVSIGAVYFENNGTYGLAFWGTSLINGEPGYTIGVLAFFSVIHSSYICGAAPHEE 285  
Db 333 GFMIGAGELMYSNYGLGFMWGSRLVDG--AVDGDILTVLMAILIGSFSLGNVSPNAQ 390  
QY 286 TFAIARGAFFHPOVIDOKPSIDNFSAGYRPESIEGVEKKNVSPNPSRKILKGL 345  
Db 391 AFTMAVAAAANKFTIDRQSPLDPSYNEGKTLDFHEGHIELRNKHYPSRPEVTYMEDV 450

QY 346 NLRKSGEVALVGLNGSGSTVVOILLQRLYPDDEGFIWDENDIRALNVRHRIHGVV 405  
Db 451 SLSPAKKTALVGPSSGSKSTVGLVERFYMPVGTVLDDGHIDKIDLNRLRQOISLV 510  
QY 406 SQEPVLEGTITISNNIKG-----RDVTDDEMRARENAVFIEMFNKFTIV 456  
Db 511 SQEPVLEGTITIKRIRGLIGTKYENESDKVRELIENAKMANAHFTALPGEFTNV 570

QY 457 GEGQAGSGOKORIAIARALVNRPKILLDEATSAIDSEKSAVOALEK 507  
Db 571 GQGFLLSGGOKORIAIARAVSDPKILLDEATSAIDTRKSEGVQALER 621

## RESULT 10

US-09-328-320-2  
; Sequence 2, Application US/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/328,320  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996,545  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11766  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1334 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-328-320-2

Query Match 36.1%; Score 937; DB 4; Length 1334;

Best Local Similarity 41.2%; Pred. No. 2.7e-83;

Matches 194; Conservative 96; Mismatches 169; Indels 12; Gaps 3;

QY 47 LYYVGIGVAAALIFGYIOISLITTAAROTKRIKQFHSVLAODIGMFDSCDDELNTRM 106  
Db 153 LYYVLLIGGEVTVYVSTGVGTYGEHATOKIRREYLESILRONIGYFDKIGAGEVTRI 212  
QY 107 T-DIDKISDGIKIALLFQNMSTFSGIAGLVKWKLTFTVLTSTPLIMASAAACSRM 165  
Db 213 TADTNLIQDGISEKVGTLTALATFVTAFTIAYYKMKLALICSTIVALVLTNGGSSQF 272  
QY 166 VISLTSKELSAVSKAGAVAEVLSSIRTVIAFRAOEKELQRYTONLKDADFGIKRTIAS 225

Db 273 IIKSKSLDSYGAGTVAEEVISIRNATAFGTQDLAKQVEVHLDEAEKMGTKNOIYM 332  
QY 226 KVSIGAYFFNMNGYGLAFWYGTSLIINGEGYITGVLAFFVSIHSSYTCGAAPHE 285  
Db 333 GFMGAFFGLMYSNYGFGFMGSRFLVDG--ADVGDILVLAAILGFSGLGNVSPNAQ 390  
QY 286 TFAIARAFAHIFOVIDKPSIDNESTAGYKPESECTVEFEKNVSPNPSRKILKGL 345  
Db 391 AFTNAVAALAAKIFCTIDROSPLDPSNEGKTLDHFECHIELRNKXHPSPREVTMEDV 450  
QY 346 NLRKSEFVALVGLNSGKSTVQLLQRLYDPDDGFIMVENDIRALNVRRHVDHIGV 405  
Db 451 SLSPACKTALVPGSSGSKSTVGLVERFYMPYRGVLLDGHDIKDLNLMWLKQOISLV 510  
QY 406 SQEVLFGTITSNNIKG-----RDVYDEMERAREANAYDFIMEFPKFNLT 456  
Db 511 SQEVLFGTITSNNIKG-----RDVYDEMERAREANAYDFIMEFPKFNLT 456  
QY 457 GEGKAGSGGOKRIATARALVVRPKLILDEATSALDSSEKSAVOALEK 507  
Db 571 GQRFLLSGGOKRIATARAVSDPKLILDEATSALDTKSEGVQALER 621  
RESULT 11  
US-08-395-246C-2  
Sequence 2, Application US/08395246C  
Patent No. 5773214  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
TITLE OF INVENTION: ASPERGILLUS FLAVUS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
City: Indianapolis  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/395,246C  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas G.  
REGISTRATION NUMBER: 35784  
REFERENCE/DOCKET NUMBER: X9683  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
TELEFAX: 317-277-1917  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1307 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-395-246C-2  
Query Match 35.5%; Score 920; DB 1; Length 1307;  
Best Local Similarity 37.7%; Pred. No. 1.3e-81;  
Matches 195; Conservative 109; Mismatches 195; Indels 18; Gaps 6;  
QY 6 LASLVNGACPLPLMLVLGEMSDNLISGLVOTNTYSEF-----RLTLYVVGIGVAALIFG 60  
Db 67 LAATAGGAALPLFLVFLGRILT-STFODIATHRIRIYDHFHHHLTNVYVFIYLGAAEFVAI 125

QY 61 YIOLSMITTAARQTKRIKQFPHSVLAODIGWFDSCDIGELNRMV-DIDKISDGIK 119  
Db 126 YLATVGFITYDHHVQQRVEYFOAILRONTAFEDTLGAGETTRITRADNLIJDGISEK 185  
QY 120 IALLFQNMSPESIGLAVGLVGMKLTTLVTLSTPLINASAACSRMVISLTSKELSAVSK 179  
Db 186 VGLALTLGISTFTVAFTIAYIKMKLALICSAALLALLTLMGCGSTLMIFSKKALEVOGR 245  
QY 180 AGAAVEEVLSSIRIVAFRAOEKELQRTQNLKPAKDFIKRTTASKVSIGAVFFPMNGT 239  
Db 246 GASMAEDILDSIRFVAAFNAOETLARKYESHLKDAEGGMSKVIYFALMGALLCIMLN 305  
QY 240 YGLAFWYGTSLIINGEGYITGVLAFFVSIHSSYTCGAAPHEPFAIRGAAPHIIG 299  
Db 306 YGLGFMGSRFLVAGISNIKAGDVLTTMAAILLGSYLNIGNVAPGQALSDVAALASKIYG 365  
QY 300 VIDKPSIDNESTAGYKPESECTVEFEKNVSPNPSRKILKGLNLRKSGEVALVG 359  
Db 366 TIDROSPLDALSDDGKTLEFVRGNIVLQNIHVPSPREYVAHDLSCTIPAGKTTAFVG 425  
QY 360 LNSGKSTVQLLQRLYDPDDGFIMVENDIRALNVRRHVDHIGVSOEVLFGTITSNN 419  
Db 426 PGGSKSTIILRFRYDPAVTIMLDGHDIOIOTNLRLRQMSLVSOEPLFAFTIAEN 485  
QY 420 IKYG-----RDVYDE--EMERAREANAYDFIMEFPKFNLTVEKGAQMGSGGOKR 470  
Db 486 IRYGIGSRFEKSTYEIRKVEEAARANAHDFTMALPNQYDINI--ESFSLSGGOKR 543  
QY 471 IAIARALVRNPKLILDEATSALDSSEKSAVOALEK 507  
Db 544 IAIARAIKDKKILILDEATSALDTKSEKLVQALDK 580

RESULT 12  
US-08-996-644-2  
Sequence 2, Application US/08996644  
Patent No. 5945324  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten  
TITLE OF INVENTION: Multiple Drug Resistance Gene atrc of  
TITLE OF INVENTION: Aspergillus nidulans  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
City: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,644  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11765  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1308 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-996-644-2

Query Match 29.4%; Score 762.5; DB 2; Length 1308;  
Best Local Similarity 36.4%; Pred. No. 4.5e-66;  
Matches 174; Conservative 103; Mismatches 176; Indels 25; Gaps 6;

47 LYYVGIGVALLFFGYIQLSWITTAARQTRIRKQFHSVLAQDQIGMFDSCDQIGELNTRM 106  
141 LFFVYIGIARLVSTYNNLTLYAAVRIYRNIRHAYLKALSOEVAYYDFGSGGSAQA 200  
107 TDIDK-ISDQIGDKIALLFQNNSTFSGIAGLVKWKLTLY-----TLSTPLIMAS 158  
201 TSNGLKIQAGSDKIGLIFQGLAFAVTLTSLRLMCKMKTLLICICIPVATIGTGVAAV 260  
159 AACSMMVLSLTKELSAVSKAGVAEVLSSIRVIYAFRAQEKELQRTONLKDADKG 218  
261 EAGHETRILOI-----HAQANSFPEGILAGYKAVHAGMRDLSVRKDEYVEAHKYG 313  
219 IKRTIASKVSIGAVFFPMNGTYGLAFWYGTSLILNGEPGYTIGTVLAFVFSYHSSYCIG 278  
314 KTSPLGLLFSAEYITTYGLAWOGIHMGRGEIG-TAGDIFTVLLSVIASINLT 372  
279 AAVPHETAIARGAAPHIYQVYDKKPSIDNFTAGYKPEISIGYVEFKNVSNYSRPS 338  
373 LLAPEYIESFRASAAAOFLRLIDRESEINPYKEGLEPERVLGDELEWTFSTYPTREG 432  
339 IKTLKGLNLRKSGETVALVGLNGSGKSTVVOQLQRLYDPDDGFIMWENDIRALNVRHY 398  
433 ITVLDNFSILKVPAGKYATALVGSGSGKSTIVGLERKYNFTSGAINDGLISELVNGL 492  
399 RDHIGVYSEPVLEFQTTSNNIKY-----RDDVTDEMER--AARENAVDFIMEFP 449  
493 RRVNRLVQGPVLEFGSVFQFNIRKGLVGPWENASREQMERVOEAKLAIAHEFTSELT 552  
450 NKENTLVGEKAGMSGGOKORIALARLVNPKILLDETSALDSESKANQAALX 507  
553 DGYDTLIGERGGLISGGOKORVARSVSOQPVLLDEATSLDPHAETIVQKALDK 610

RESULT 13

US-09-352-552-2

; Sequence 2, Application US/09352552  
; Patent No. 6060264  
; GENERAL INFORMATION:  
; APPLICANT: Skatrud, Paul L.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrc of  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrc of  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: U.S.  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/352,552  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996,644  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Webster, Thomas D.  
; REGISTRATION NUMBER: 39,872  
; REFERENCE/DOCKET NUMBER: X-11765

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3334  
; TELEFAX: 317-276-2763  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1308 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-352-552-2

Query Match 29.4%; Score 762.5; DB 3; Length 1308;  
Best Local Similarity 36.4%; Pred. No. 4.5e-66;  
Matches 174; Conservative 103; Mismatches 176; Indels 25; Gaps 6;

47 LYYVGIGVALLFFGYIQLSWITTAARQTRIRKQFHSVLAQDQIGMFDSCDQIGELNTRM 106  
141 LFFVYIGIARLVSTYNNLTLYAAVRIYRNIRHAYLKALSOEVAYYDFGSGGSAQA 200  
107 TDIDK-ISDQIGDKIALLFQNNSTFSGIAGLVKWKLTLY-----TLSTPLIMAS 158  
201 TSNGLKIQAGSDKIGLIFQGLAFAVTLTSLRLMCKMKTLLICICIPVATIGTGVAAV 260  
159 AACSMMVLSLTKELSAVSKAGVAEVLSSIRVIYAFRAQEKELQRTONLKDADKG 218  
261 EAGHETRILOI-----HAQANSFPEGILAGYKAVHAGMRDLSVRKDEYVEAHKYG 313  
219 IKRTIASKVSIGAVFFPMNGTYGLAFWYGTSLILNGEPGYTIGTVLAFVFSYHSSYCIG 278  
314 KTSPLGLLFSAEYITTYGLAWOGIHMGRGEIG-TAGDIFTVLLSVIASINLT 372  
279 AAVPHETAIARGAAPHIYQVYDKKPSIDNFTAGYKPEISIGYVEFKNVSNYSRPS 338  
373 LLAPEYIESFRASAAAOFLRLIDRESEINPYKEGLEPERVLGDELEWTFSTYPTREG 432  
339 IKTLKGLNLRKSGETVALVGLNGSGKSTVVOQLQRLYDPDDGFIMWENDIRALNVRHY 398  
433 ITVLDNFSILKVPAGKYATALVGSGSGKSTIVGLERKYNFTSGAINDGLISELVNGL 492  
399 RDHIGVYSEPVLEFQTTSNNIKY-----RDDVTDEMER--AARENAVDFIMEFP 449  
493 RRVNRLVQGPVLEFGSVFQFNIRKGLVGPWENASREQMERVOEAKLAIAHEFTSELT 552  
450 NKENTLVGEKAGMSGGOKORIALARLVNPKILLDETSALDSESKANQAALX 507  
553 DGYDTLIGERGGLISGGOKORVARSVSOQPVLLDEATSLDPHAETIVQKALDK 610

RESULT 14

US-08-394-880B-2

; Sequence 2, Application US/08394880B  
; Patent No. 5705352  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene Of  
; TITLE OF INVENTION: Aspergillus fumigatus  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Lilly and Company/Patent Division  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: US  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,880B

FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant G., Thomas  
REGISTRATION NUMBER: 35784  
REFERENCE/DOCKET NUMBER: X-9682  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (317) 276-2459  
TELEFAX: (317) 277-1917  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-394-880B-2

Query Match 25.8%; Score 669.5; DB 1; Length 791;  
Best Local Similarity 32.6%; Pred. No. 3e-57;  
Matches 169; Conservative 99; Mismatches 222; Indels 29; Gaps 10;

1 MILGLIATLVNGACLPMPVLYGEMSDNLISGLVQNTYTSFFRLTYL-----VGIQV 54  
181 LALAFLLVSSGIMTSIPFSIGKIMDTSTATTEGNEFLGLSLPMFYGALAGLITLIGA 240  
55 AALITGYIOLISMITTAROTKRIKOFPHSVLAODIGMFDSCDIGELNTRMTD----ID 110  
241 AA--NYRIITLITVGERIVARLSKLEFQTEVDAEFEDANRQGDILSRSSDTIIVG 297  
111 K-ISDIDGDKIALLFQNMSTFSGIAGLVGKMKLTLYLTSLSPILMASAACSRLVSL 169  
298 KSTIONLSDGLRAVNSGAGFGLMAYSL---KLSSILALLPITIGLAFYGAIRNL 353  
170 T---SKELSAVSKAGAAVEVLSTIRIVIAFRAQEKELQRTQNLKADKFGIKRTIASK 226  
354 SROIQRNIGTGLTK---IAEERLGNVKTQSOFAGEVLEVRVYNNQVRKIFELGKRESLISA 410  
227 VSLGAVYFFMNGTYGLAWYGTSLILNGEPGTYTIGTVALVFFSVIHSSYCICAAAPHET 286  
411 TFFSSTGRAGNTTILALLYGGGAVQSG--ATTIGELTSLMYTAVASSMSGLSSFFSE 468  
287 FAIARGAFFHIFOVIDKKPSIDNFSTAGYKPSIEGVEEKNVSPNPSRSPIKILKGLN 346  
469 LMKGGAASRLFELODRPTIS--PTKGEKVASANGPIREFEVTTSYTPRPVAPLPIFDLN 526  
347 LRKSGETVALVGLNGSKSTVQVLLQRLYPDDEGFIWVENDIRALNVRYRHHIGVYS 406  
527 FEIPQGTWVAIVGSPGGKSTIASILIRFYSPTGRLVIGKDIYHMAKSLRRKIGIVS 586  
407 QEPVLEFETTSNNIKYGGDDVTDEMEARAARANAIDFIMEPPKFNLLYGEKGQOMSGG 466  
587 QEPVLEFETTSNNIKYGGDDVTDEMEARAARANAIDFIMEPPKFNLLYGEKGQOMSGG 645  
467 QKORIAIARALVRNPKLILDEATSAIDSESKSAVOAAL 505  
646 QKORIAIARALVRNPKLILDEATSAIDSESKSAVOAAL 684

RESULT 15  
US-08-232-537-2  
Sequence 2, Application US/08232537  
Patent No. 5516655  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis

STATE: Indiana  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232.537  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas G.  
REGISTRATION NUMBER: 35,784  
REFERENCE/DOCKET NUMBER: X9212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
TELEFAX: 317-276-1917  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1302 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-232-537-2

Query Match 24.7%; Score 639.5; DB 1; Length 1302;  
Best Local Similarity 32.0%; Pred. No. 6.2e-54;  
Matches 173; Conservative 108; Mismatches 203; Indels 57; Gaps 14;

3 LGIATLVNGACLPMPVLYGEMSDNLISGLVQNTYTSFFRLTYL-----VGIQV 48  
744 LGIATLVNGACLPMPVLYGEMSDNLISGLVQNTYTSFFRLTYL-----VGIQV 48  
49 --YVIGIYGAALFFGYIOLISMITTAROTKRIKOFPHSVLAODIGMFDSCDIGELNTRM 106  
778 FSYAGNCC---FG-----IVSHYAKIQLHSLASILRQDMQWISGQVPSLMSSL 826  
107 -TDIDKISDIDGDKIALLFQNMSTFSGIAGLVGKMKLTLYLTSLSPILMASAACSRLV 165  
827 SSDAQQLACLSGVALIGTITFTVCVSTIGGIIAHVAMKIAVLLAAPPMTAGVRLRV 886  
166 VYLSLTKELSAVSKAGAAVEVLSTIRIVIAFRAQEKELQRTQNLKADKFGIKRTIASK 224  
887 IALAESRHRSAVNDAASTIAEACRGIRT--IASLGRERGVSRAVNAVEPYDKGIRFTLI 945  
225 SKVSLGAVYFFMNGTYGLAWYGTSLILNGEPGTYTIGTVALVFFSVIHSSYCICAAAPHET 284  
946 TMTLALASISITYFYVALAYVMGAKOVNRGT--YSQLODFITVLPALFLPSAQSAGQIFSLIS 1003  
285 ETFAIRGAFFHIFOVIDKKPSIDNFST--AGYKPSI-----EGTVEERN 328  
1004 PEMASAGVAAARVFGHIDQKPTIVDVAKQSGALPSTSLTIPTLBDKASPSGGMIEKN 1063  
329 VSFNTPSRPSIKILKGLNIRKSGETVALVGLNGSKSTVQVLLQRLYPDDEGFIWVENDIR 388  
1064 VSLCYPSKRPQHPALONVNISIRPGEFIALVGPAGKSTILSLDRFPTAGSVLDQ 1123  
389 DIRALNVHRHYRDHIGVSEPVLFETTSNNIKYGRD--VTDEMEARAARANAIDFIMEPPK 445  
1124 DIREVAFQHRKRLGLVQEPFLPFGSISYNGIGLAAEQGLVTRDIEKICKACGIIHEFI 1183  
446 MEFPNKTVLVEKGAGQSGGQKORIAIARALVRNPKLILDEATSAIDSESKSAVOAAL 505  
1184 MSLPFGYSPGECTGNSKLSGGQKORIAIARALVRNPKLILDEATSAIDSESKSAVOAAL 1243  
506 E 506  
1244 D 1244



Tue Apr 23 09:35:41 2002

us-09-873-409-8.ra1

Page 10

Search completed: April 22, 2002, 14:35:04  
Job time: 136 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 15:08:27 ; Search time 9392.76 Seconds

(without alignments)  
2363.605 Million cell updates/sec

Title: US-09-873-409-9

Perfect score: 2066  
Sequence: 1 cgaagcaagtcgtgactata.....atgcacagtcagtcagttga 2066

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_num:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rpd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	386.6	18.7	943	10	AL520322
2	275.2	13.3	726	11	BG293345
3	273.2	13.2	1019	11	BG248052
4	269	13.0	944	11	BF796582
5	262.2	12.7	785	10	AV709991
6	252.4	12.2	560	10	BE749379
7	249.4	12.1	981	11	BF969667
8	249	12.1	795	11	BF313560
9	242.6	11.7	780	11	BG587938
10	232.4	11.2	886	13	BH139685
11	229.4	11.1	482	11	BI344244
12	227.8	11.0	651	10	BE016688

13	224.6	10.9	947	13	AZ683753	AZ683753	ENTL96TF
14	221.4	10.7	886	13	AZ540627	AZ540627	ENTEO18TF
15	220.4	10.7	897	13	AZ541090	AZ541090	ENTDS67TR
16	219.8	10.6	834	13	AZ548312	AZ548312	ENTFG07TR
17	218	10.6	880	13	AZ687805	AZ687805	ENTLUS2TF
18	217.6	10.5	871	13	AZ682350	AZ682350	ENTRBL6TF
19	215.4	10.4	932	13	AZ670821	AZ670821	ENTJN69TF
20	213	10.3	519	10	AA434959	AA434959	ve18c12.r
21	213	10.3	853	13	AZ679807	AZ679807	ENTH167TR
22	212.2	10.3	605	10	AV616675	AV616675	AV616675
23	211	10.2	823	13	AZ532602	AZ532602	ENTCR50TR
24	209.6	10.1	891	13	AZ682250	AZ682250	ENTRT68TR
25	209.2	10.1	919	13	AZ690701	AZ690701	ENTRT58TR
26	208	10.1	851	11	BF969062	BF969062	60226998TR
27	206.4	10.0	687	11	BG646725	BG646725	EST508344
28	205.4	9.9	899	13	AZ548799	AZ548799	ENTDK47TR
29	202.6	9.8	921	13	AZ687628	AZ687628	ENTIR96TR
30	200.4	9.7	787	11	BG584063	BG584063	EST485823
31	200	9.7	842	13	AZ671925	AZ671925	ENTMN93TR
32	197	9.5	541	10	A1722858	A1722858	fc32f01.y
33	192.6	9.3	511	10	AA239727	AA239727	mx81b05.r
34	191	9.2	1201	11	BF304387	BF304387	601887302
35	189.4	9.2	392	10	A1115290	A1115290	ub87h07.r
36	188.8	9.1	822	11	BI332761	BI332761	602979574
37	187.6	9.1	759	10	BE376459	BE376459	601226251
38	187.4	9.1	616	10	A1728473	A1728473	BNGH1108
39	187	9.1	388	11	BF190122	BF190122	236402.MA
40	187	9.1	551	11	BG512708	BG512708	dad28404
41	185.6	9.0	680	11	BG646029	BG646029	EST507648
42	183.8	8.9	854	13	AZ682568	AZ682568	ENTMWS5TF
43	182.8	8.8	760	10	AL520321	AL520321	AL520321
44	180.8	8.8	886	13	AZ668396	AZ668396	ENTM61TF
45	180	8.7	397	11	BG875577	BG875577	IL5-CN002

#### ALIGNMENTS

RESULT 1  
AL520322  
LOCUS AL520322 943 bp mRNA  
DEFINITION AL520322 LTI\_NFL004.NBC2 Homo sapiens cDNA clone CS0DB006YC15 5  
prime, mRNA sequence.  
ACCESSION AL520322  
VERSION AL520322.1 GI:12783815  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 943)  
REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
AUTHORS Full-length cDNA libraries and normalization  
TITLE Unpublished (2001)  
JOURNAL  
COMMENT  
CONTACT: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source  
1..943  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DB006YC15"  
/clone\_1lb="LTI\_NFL004\_NBC2"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: brain. Vector: pCMVSPORT 6: 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life

QY 1729 acacacaagtcgacitgaagsgcaccagcttctctgcggtccagaacaacagcctcaca 1730

	Query Match	Score	DB 11	Length	1019;
Best Local Similarity	63.2%	Pred No.	1.6e-58		
Matches	437;	Conservative	0;	Mismatches	253;
				Indels	2;
				Gaps	1
Qy	1326	gaagggaattcagagcttcgcagaagctctctcttcctatccatgctgcgcgaagatgtttc	1385		
Db	4	GAAGGAATGTGTCACATTTAGTGGAGTCGCTGATTCACATTCGCCACCGACACATTCGCCA	63		

```
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM955 row: n column: 15
High quality sequence start: 9
High quality sequence stop: 669
```

Location/Qualifiers  
1. .944

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4341710"
/clone_id="NH_MCC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1,867 kb. Constructed by Life Technologies full-length clones and library by Life Technologies Note: this is a NH_MCC library."

```

BASE COUNT	265 a	241 c	237 g	201 t
ORIGIN				

Query Match	Score	DB	Length
13.08;	269;	11;	944;

Matches 409; Conservative 0; Mismatches 215; Indels 1; Gaps 1;

**DY**    1224 tattccaagccaaatcgggggtcgcgatctgttgccttttgtaagaaccaa  
         ||| |||| |     |     ||| |||| ||  
**Db**        39 TATGCTAAAGCTTAACGTGTTCGACCACATTATCATGCGTTTGAAGAACAACCTCTG 98

QY 1284 atagacagccgcagttcaagaagggaaaaagccagacacatgcygaaggaattttagagttc 1343

Db 99 atttcacagctacacatgaaagagggcgctgaagcctgatnaatttgaagcaaatatatacatattt 158

QY 1344 cgagaagtcctccttctctatccatgcgcccagayttttcaacccctcggtgatcc 1403  
||||| ||| ||||| | ||| ||| |  
Db 159 AATGAAGTCGTGTTCAACTATCCACCCGAGACAACGTCGCTTACGGCCTGAGC 218

```
QY   1404 cccagtatcyagcgaggaaayaacagctacgattgttggggagcgcgcgtcttgysaaaagc 146
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   219 CTGGAGGTGAAGAAGGCCACACACTAAGCCCTGGGTGGGCAGCAGTGGCTGTGGAAGAGC 278
```

Oy 1464 actctgttcaactcttcgcagagaactttatgaaccccgatgacaagtagctgtttgat 1522  
|| || || || || || || || || || || || || || || || || || || ||  
Db 279 ACGGAGGTCCAGCTCTGGAGACGGTTCACGACCCTTTGGCGGGGACACTGCCTTCGAT 338

Qy 1524 ggctgtgatacaaaagaattgaatgtacagtggtctcgcttcccaatagcaatggtcct 1583  
||| || ||||| ||| ||||| ||||| ||| ||||| ||  
Db 339 GGTCAGAACCAAGCAACTCAATGTCACAGTGGCTCAGAGCTCAACTCGGAATGTTCT 398

QY 1584 caagagccctgctctcaactgcgcatctgtgagacatcgctaigtgacaacagc 164  
|| ||||| | | | | | ||||| | | | | | ||||| | | | | | |||||  
Db 399 CAGGAGCCCTATTCCTATTGACTGCAGCATTCGCCGAGATATTGCCCTAGGAGACAACAGC 458

QY 1644 cgtgctgctccattagatgagatcaagaagcgcgaatgcagcaatafatcatctttt 170  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 459 CGGGTTGTATCACAGGATGAATTCTGTAAGTGACGCCAAGCTGCACAATATACCTTTC 518

QY 1704 attgaagtcctccctgagaataacaaacacaagtgtgactgaagpacacagcttctc 176  
||| || | | | | | | | | | | | | | | | |  
Db 519 ATCGAGACGTTATCCCAACAATATGAACAACAGTGGGAGATAAGGGGACTCAGCTCTCA 578

QY 1764 ggcgcgcagaacaagaactgctatgtcgaaggctcttccaaaacc-aaaattt 182

QY 1823 atgttgatgagccactcagcc 1847  
|||||  
ph 639 CCTGTTGGAGTAAAGCTACATCAGCC 663

RESULT 5  
AV700001

LOCUS	AV709991	785 bp	mRNA	EST	09-OCT-2000
DEFINITION	AV709991	Cu Homo sapiens cDNA clone CUAAVA09.5',	mRNA sequence		
ACCESSION	AV709991				
VERSION	AV700001	1	01-10720373		

KEYWORDS EST.

SOURCE	human
ORGANISM	Homo

SOURCE	ORGANISM
human.	<i>Homo sapiens</i>
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 785)
AUTHORS	Peng, Y., Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N.,

**TITLE**  
Homo sapiens cDNA Cu clones  
**JOURNAL**  
Unpublished (2000)  
**COMMENT**  
Contact: Zequang Han

Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzygc@cnhg.sh.cn  
This clone is available at CHGC in Shanghai.

LEAD SOURCE

```

/db_xref="taxon:9606"
/clone="CUAAJA09"
/clone_lib="Cu"

```

```
syndrome"  
/dev_stage="Adult"  
/note="Vector: pbluescript sk(-)"
```

ORIGIN

Best Local Similarity 64.0%; Pred. No. 9.5e-56;  
Matches 458; Conservative 0; Mismatches 254; Indels 4; Gaps

Db 73 GGAATTGGAMGAGTATTTTCAGCTGTTGCTTTGGTGCATGGCCGTGGGGCAAGTCAGT 132

Db 133 TCATTGGCTCTGACATATGCCAAAGCCAAATATCAGCAGCCACATCATCATGATCATTT 192

Db 193 GAAAAAACCCCTTTGATTGCACCTACAGCAGCGAAGCCTAATGCCGAACATTTGGAA 252

[illegible][illegible]

Dy 1449 ggcctggggaagcactctgttcaactcttcgagagacattatgaccctcgcaaggaa 1508  
| | | | | | | | | | | | | | | | | |  
Db 373 GCGCTGGGAAGCACA GTGCTCCAGCTTC TGAAGCGGTTCACGACCCTTGGCAAGG 432

Qy 1509 caagtcgtcttgatgigtgtgatgcacaagaattgaatgacagtggtccgccca 1508  
||||| ||||| || ||||| ||||| |||||  
Db 433 AAGTCTCTGTTGATGGCAAGAATTAAAGCAGACTGAATGTTACAGTGGCTCCGAGCACAC 492

Qy 1569 atagcaatcgttccctcaagaagcctgtgccttcacactgaagcatgtctgagaacaatgcgc 1628  
| | | | | | | | | | | | | | | | | |  
Db 493 CTGGCATCGTCGCCAGAGAGCCCATCTGTTGACTGCACATTGCTGAGAACAATTGCC 552

Qy 1629 tatggtcaacaacagcc-gtgtgtgtccattagatgagatcaagaagccgaatgtcagc 1687  
|||||  
Db 553 TATGAGACAACAGCCGGGGTGTGTACAGAGAGATTGTGAGG66CAGCAAAAGGA66C 612

QY	DB	1595	gctcttaactgacgcatctgtgagacacatgcgtctatggtgacaacagccgltgtgtgc	1654
QY	1688	aaatccatcttcttattgaaagcctccctgaggaataacacacaaagttgagctgaa	1747	
DB	613	CAACATACATGCTTCATCAGACTCAGTGCTT	ATTAATATATAGCACTTAACACTAANGAGACAA	671
QY	1748	aggagacaagcttcttgccgcgcagaacaacagactagctatctgaaggctctctcca	1807	
DB	672	AGGACCTCAGCTCTCTGGGGGCCACAGAAACACAA	TTGGCATTAGCTTGTGGCCCTGTGA	730
QY	1808	aaaccacaaatttattgttgatgagggccaacttcagccctcgataatgacagt	1863	
DB	731	CAGACTTATAA	TTTGTTTTGGATGAAGCCACGTCCACTCTTGATACAGAAATG	785
RESULT	6			
BE749379	LOCUS	BE749379	560 bp	mRNA
ACCESSION	DEFINITION	200194 MARC	4BOV Bos taurus	CDNA 5', mRNA sequence.
VERSION	KEYWORDS	BE749379	1	GI:1016371
SOURCE	ORGANISM	COW	Bos taurus	
REFERENCE	AUTHORS	Smith,T.P.L., Grosse,W.M., Fekling,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,D.E., White,J., Cho,D., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,M.W., Rohrer,G.A., Chitko-McKown,C.G., Keeler,J.W., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and		
TITLE	JOURNAL	Sequence evaluation of four pooled-tissue normalized bovine CDNA		
MEDLINE	COMMENT	libraries and construction of a gene index for cattle		
CONTACT	CONTACT	Smith TPL		
PO BOX	PO BOX	166		
CLAY CENTER	CLAY CENTER	NE 68933-0166		
USA	USA			
TEL	TEL	402 762 4366		
FAX	FAX	402 762 4390		
EMAIL	EMAIL	smith@email.marc.usda.gov		
SINGLE PASS SEQUENCING	SINGLE PASS SEQUENCING	Bases called and alt_trimmed with phred		
V0.980904.e	V0.980904.e	Vector identified by cross_match with the -minscore 18		
-minmatch 12 options	-minmatch 12 options			
PCR PRIMERS	PCR PRIMERS			
FORWARD: AGGAACAGCTATGACCAT	FORWARD: AGGAACAGCTATGACCAT			
BACKWARD: GTTTCCACATCAGCAGC	BACKWARD: GTTTCCACATCAGCAGC			
PLATE: 108	PLATE: 108	ROW: F COLUMN: 11		
SEQ PRIMER: ATTTAGTGACATATTA	SEQ PRIMER: ATTTAGTGACATATTA			
LOCATION/QUALIFIERS	LOCATION/QUALIFIERS			
1..560	1..560			
/organism="Bos taurus"	/organism="Bos taurus"			
/db_xref="taxon:9913"	/db_xref="taxon:9913"			
/clone_lib="MARC 4BOV"	/clone_lib="MARC 4BOV"			
/library_type="pooled"	/library_type="pooled"			
/lab_host="DH10B"	/lab_host="DH10B"			
/note="Vector: pcMV SPORT6; Site 1: XbaI; Site 2: XhoI;	/note="Vector: pcMV SPORT6; Site 1: XbaI; Site 2: XhoI;			
library made from pooled tissue from day 20 and day 40	library made from pooled tissue from day 20 and day 40			
embryos."	embryos."			
BASE COUNT	BASE COUNT	154 a	153 c	146 g
ORIGIN	ORIGIN			107 t
Query Match	Query Match	12.2%	Score 252.4	DB 10; Length 560;
Best Local Similarity	Best Local Similarity	67.5%	Pred. No. 2.7e-53;	
Matches 355; Conservative 0; Mismatches 171; Indels 0; Gaps 0;	Matches 355; Conservative 0; Mismatches 171; Indels 0; Gaps 0;			
QY	1535	aaaagaatgaaatgacagtgagctcgttcctcccaataagaacatcgttcctcaagagcgtgt	1594	
DB	1	AAAGAAGCTCAACGTCGACGTCCTCAGAGCCCAACTTGGAAATCGTGTGACAGACCCGT	60	
QY	1595	gctcttaactgacgcatctgtgagacacatgcgtctatggtgacaacagccgltgtgtgc	1654	

FEATURES	source
BASE COUNT	267 a 194 c 250 g 270 t
ORIGIN	
Db	61 CCGTTTACAGTCAGCACCATTGGCCGACACAAATTTGCCATTATG666GACAACAGCGGCGCTGTAAc 120
Qy	1655 attagatgagatcaaaagaagccgcaaatgacgaaataatccattcttatttgaaggtct 1714
Db	121 CATGCTGTAATTTGTGAGCGCGCAGCCCAAGCCAAACATTCATCCTTTCATTGAGACCTT 180
Qy	1715 cccftgagaataacacacaaagtttgagactgaagaagacaaagcttcttcgcgccaa 1774
Db	181 GCCCCACAAATTTGTAACAAAGAGTGTGGAGATTAAAGGGGACTCAGCTTTCGGGGGACAGAA 240
Qy	1775 acaagactagactatgtgcaaggcgctctctccaaaacccaaatttattgtgtatga 1834
Db	241 ACAGAGGATGTGATTATGGCCGAGGCCCTTATCCGACACCCCGCATCTACTGCTGGATGA 300
Qy	1835 ggcacattcagccctcgaatagtacagtgagaaggttggttcagcatgcccctgtataaagc 1894
Db	301 AGCCACGTCAGACACTGGAATCTGAAAGTGAGAAATTTGCCAAGAACCCCTTGACAAAGC 360
Qy	1895 caggaacggaagacatgctcagtggtgtcactcaagctctctgtcaattcagaacgaga 1954
Db	361 CCGAACAAGCGCCGACCTCATCTGATGCTGCTCACCGGCTGTCCACATCCAGAAAGCAGAGA 420
Qy	1955 ttgtgtagtggttcgcgcaatggaagaataaagaagaacgaactcaagaagctcct 2014
Db	421 CTGTGATGATGTTGATTTGCAATGGCAGAGTCAGGAGCAGCAGGACACACCCAGCAGCTGCT 480
Qy	2015 gagaatcgaagacatatatttgaagttagttgaatgcacagtcagtg 2060
Db	481 GGCACAGAAAGGCATCTATTTCACCATGTGTCAGTGTCCAGGCTGGG 526
RESULT 7	
BP696667	981 bp mRNA EST 22-JAN-2001
LOCUS	6022272046p1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4360090 5,
DEFINITION	mRNA sequence.
ACCESSION	BF696667
VERSION	BF696667.1 GI:12336882
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE	1 (bases 1 to 981)
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@b-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L1AM1000 row: 1 column: 11 High quality sequence stop: 670. Location/Qualifiers 1. 981 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4360090" /clone_lib="NIH_MGC_84" /tissue_type="adrenal cortex carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site:1; Note: Site:2: Salt; Cloned unidirectionally; oligo-dT primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."



Db 652 TGTCCACCATTCAGATAGACTT--ATAGTGTGTTTACAGATGACAGCTCAGCGG 709

Qy 1993 aaggacatcatcaagagctcctgagaatcgacatatatttaagtag 2043

Db 710 CATGGACGCATCAGAGTCTGTGACAGAAAGCCTATTATTTCACGCTAG 760

## RESULT 9

BG587938

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 780

/organism="Medicago truncatula/Glomus versiforme mixed EST

/library="

/cultivar="Medicago truncatula genotype A17"

/db\_xref="taxon:119092"

/clone="PMHAM-51L19"

/clone\_lib="MHAM"

/tissue\_type="roots colonized with Glomus versiforme"

/dev\_stage="roots harvested at 10, 17, 22, 31 and 38 days

post-inoculation with Glomus versiforme. The library was

made from a mixture of RNA from each of these stages."

/lab\_host="E. coli strain XL0LR"

/note="vector: pBluescript SK-; Site\_1: EORI; Site\_2:

XhoI; cDNA was prepared from polyA+ enriched RNA from

roots harvested at 10, 17, 22, 31 and 38 days

post-inoculation with Glomus versiforme. The cDNA was

directionally ligated into the UniZap XR vector from

Stratagene and packaged using Gigapack III Gold packaging

extracts. Plasmids containing cDNA inserts were excised

from the recombinant lambda-Zap phage using Ex-assist

helper phage and propagated in XL0LR cells."

BASE COUNT

ORIGIN

234 a

156 c

184 g

206 t

Query Match

Best Local Similarity

Matches 444; Conservative

0; Mismatches 309; Indels 6; Gaps 1;

Qy 1219

cgaatattccaagaacgaggggctgcatctgttgcctgttggaagaagaac 1278

Db 28

CTGCAATTTCAAAAGCTTAAGTTGCGGCTGCTTAAGATTTCGGAATTAATGATCACCAGC 87

Qy 1279

caaatatagacagcgagtcgaagaagaagaagccagacacatgtgaagggaatttag 1338

Db 88

CTGGTATGATAGTAAGACAGTGAATCTGGAATTGAGACAGTACTGAGCTGTGTT 147

Qy 1339 agttcagagaagctctcttctatccatgctgcccagatgcttctatccctcgtagc 1398

Db 148 AACTGAAAATGTGACCTTCTTATCCATCAAGACCTGAAGTTCTGATCTCAATGATT 207

Qy 1399 taccctcagatctgagcgaagaagaagtagcatctgttgagggaagcagcgtctgtgga 1458

Db 208 TCTCCTTGAGTGTCTCCGCGGGAAGACCATACCTTTAGTTGAGACAGGCGCTGTGCA 267

## RESULT 10

BH139685

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. 886

/organism="Entamoeba histolytica Sheared DNA Entamoeba histolytica

genomic, DNA sequence."

/db\_xref="taxon:119092"

/clone="PMHAM-51L19"

/clone\_lib="MHAM"

/tissue\_type="roots colonized with Glomus versiforme"

/dev\_stage="roots harvested at 10, 17, 22, 31 and 38 days

post-inoculation with Glomus versiforme. The library was

made from a mixture of RNA from each of these stages."

/lab\_host="E. coli strain XL0LR"

/note="vector: pBluescript SK-; Site\_1: EORI; Site\_2:

XhoI; cDNA was prepared from polyA+ enriched RNA from

roots harvested at 10, 17, 22, 31 and 38 days

post-inoculation with Glomus versiforme. The cDNA was

directionally ligated into the UniZap XR vector from

Stratagene and packaged using Gigapack III Gold packaging

extracts. Plasmids containing cDNA inserts were excised

from the recombinant lambda-Zap phage using Ex-assist

helper phage and propagated in XL0LR cells."

BASE COUNT

ORIGIN

234 a

156 c

184 g

206 t

Query Match

Best Local Similarity

Matches 444; Conservative

0; Mismatches 309; Indels 6; Gaps 1;

Qy 1219

cgaatattccaagaacgaggggctgcatctgttgcctgttggaagaagaac 1278

Db 28

CTGCAATTTCAAAAGCTTAAGTTGCGGCTGCTTAAGATTTCGGAATTAATGATCACCAGC 87

Qy 1279

caaatatagacagcgagtcgaagaagaagaagccagacacatgtgaagggaatttag 1338

Db 88

CTGGTATGATAGTAAGACAGTGAATCTGGAATTGAGACAGTACTGAGCTGTGTT 147

Qy 1219

cgaatattccaagaacgaggggctgcatctgttgcctgttggaagaagaac 1278

Db 28

CTGCAATTTCAAAAGCTTAAGTTGCGGCTGCTTAAGATTTCGGAATTAATGATCACCAGC 87

Qy 1279

caaatatagacagcgagtcgaagaagaagaagccagacacatgtgaagggaatttag 1338

Db 88

CTGGTATGATAGTAAGACAGTGAATCTGGAATTGAGACAGTACTGAGCTGTGTT 147

Qy 1219

cgaatattccaagaacgaggggctgcatctgttgcctgttggaagaagaac 1278

Db 28

CTGCAATTTCAAAAGCTTAAGTTGCGGCTGCTTAAGATTTCGGAATTAATGATCACCAGC 87

Qy 1279

caaatatagacagcgagtcgaagaagaagaagccagacacatgtgaagggaatttag 1338

Db 88

CTGGTATGATAGTAAGACAGTGAATCTGGAATTGAGACAGTACTGAGCTGTGTT 147





QY	1652	gccatttagatgacccaagaagccgcgaatcagcaaatatccattctttatggaag	1711
Db	268	GTCCACGAGAGGAGNGATCTGCGACGGCGCCAGAGGCCACATCTCCACCCCTTATGTGAGAC	327
QY	1712	tctccctgagaataacaacacacaagcttggatctgaaaggaagcagcttctgcgcca	1771
Db	328	ACTGCGCTATTAAATGTAAACACACAGAGTGGAGACAAAGAAACGACGTCTTGGCGGCA	387
QY	1772	gaataaagacagctatctgcaagggcctcttccaaaacccaataattttctgttga	1831
Db	388	GAAGACGGCATGTGAATAGCTGTGCGCCCTGTGCAGAGCGGCTCGTATCTGCTTGTGA	447
QY	1832	tgaagccactcaagccctcgataatagacatgaga	1866
Db	448	TGAAGCTACATCGGCTGTGAGTACCAAGTAAGTAA	482
RESULT	12		
LOCUS	BE016688		
DEFINITION	BE016688	651 bp	EST
ACCESSION	BE016688		
VERSION	BE016688.1		
KEYWORDS	EST		
SOURCE	zebrafish.		
ORGANISM	Danio rerio.		
REFERENCE			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
SOURCE			
BASE COUNT	190 a	157 c	158 g
ORIGIN			

Query Match	11.0%	Score 227.8	DB 10	Length 651
Best Local Similarity	59.0%	Pred. No. 4.9e-47		
Matches 382	Conservative 0	Mismatches 265	Indels 0	Gaps 0
QY 1212	ttggctccctgaatattcccaagaagccaatcggggctgcgcattctgttgccttctgttgaa	1271		
Db 4	TTGGCTCCTGATTTTGTCTAAAGCCAAAGCTGCAGCGATGATCATTCACCAACTGTGGAG	63		
QY 1272	aagaacccaatatatgacagccgcagttcaagaaggaaaaagccagacacatgttgaagg	1331		
Db 64	AAGAACCACGAGATTGAAATATACATGATGCTGGAGAGACCTTANACCTTTTGGGG	123		
QY 1332	aatttagagtttcggaagatcctctctctctatccatctgcgccagatgttttccatc	1391		
Db 124	AACATTGATTTTCAAGGATGTGCAGTTTCTATACCCGACTCGGCCGAATGTGAAGTGCTG	183		
QY 1392	ctgtgcttaccctcagatattagcgaggaagaagacagtagcatctgttgggaagcagcg	1451		
Db 184	CAAGGTCTGAATGTGTGGTGGTCGTCAGGATCAGACTCTGGCTGTGGGACACAGCGG	243		
QY 1452	tgttggaagaagcacttctgttcaactctgcgcagagactttagacccgtgcagaagaa	1511		
Db 244	TGTGGAAGAGACACCATATTTCAGCTTTCGGAGCCGTTTTCACACCTCGCGAGGCCAA	303		
QY 1512	gtgcgtttttagtgcgttgatgacgaagaagattgaaatgtaacgtgcgttcgttcccaata	1571		
Db 304	GTGCTGTGATGACATGATCATGACAGCAAAAGTGTGAATCTGCGATGCTGCCACACAAATG	363		
QY 1572	gcaatcgttctcctcaagaagccttgcctcctcaactgcagacattgctgagaacatcgccat	1631		
Db 364	GGGGTGTGTCTCAAGAGCCCAATCTGTTCAACTGCACCATCTCTTAAGAAACATCCANTAC	423		
QY 1632	ggtgcacaagccgctgtgtgtgcattagttgagatcaagaagccgcaatgtgaagaaat	1691		
Db 424	AGAGACCAACANTCGACATCTACACAAAGAAAGATTGAAGAGGCTCCGAANNAAGCCAAC	483		
QY 1692	atccattcttatttgaagctccctcgtggaatacaacacacaagtgtgactgaagaa	1751		
Db 484	ATACCAACTTATTTCTCAACCTTCCAGATTAATATTAACACTCNTATGGAGATTAAGAAG	543		
QY 1752	gcacagcttctgcgcgcgaagaacaagaactagctattgcaaggctccttcccaaaa	1811		
Db 544	ACTCGACTAGCTGGAAGACAGAAACAAAGATCGCATGCTCGACTGTAGTGCAGAAA	603		
QY 1812	cccaaatatttattgttggatgagccacttcaagccctcgataatga	1858		
Db 604	CGCANACTCTGTCTACTCGACGAAGCGCTGCTGTATACAGA	650		
RESULT 13				
LOCUS	AZ683753	947 bp	DNA	GSS 14-DEC-2000
DEFINITION	ENT1167E Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.			
ACCESSION	AZ683753			
VERSIONS	AZ683753.1	GI:11820899		
KEYWORDS	GSS.			
SOURCE	Entamoeba histolytica.			
ORGANISM	Entamoeba histolytica			
REFERENCE	1 (bases 1 to 947)			
AUTHORS	Loftus, B., Van Aken, S. and Fraser, C.			
TITLE	Determination of clone end sequences from Entamoeba histolytica			
JOURNAL	HMI:IMS sheared DNA library			
COMMENT	Unpublished (2000)			
	Contact: Brendan J Loftus			
	Department of Eukaryotic Genomics			
	The Institute for Genomic Research			
	9712 Medical Center Dr., Rockville, MD 20850, USA			
	Tel: 301 838 0200			
	Fax: 301 838 0208			
	Email: entae@tigr.org			

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library  
 Seq primer: M13-Forward  
 Class: shotgun  
 High quality sequence start: 40  
 High quality sequence stop: 778.  
 Location/Qualifiers

## FEATURES

Source

```

1. 947
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db.xref="taxon:5759"
/clone.lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOST1; Site_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaundin and B. Barrell, Oxford University Press, 1999)."
```

BASE COUNT 379 a 121 c 178 g 269 t

ORIGIN

Query Match 10.9%; Score 224.6; DB 13; Length 947;  
 Best Local Similarity 55.8%; Pred. No. 3.3e-46;  
 Matches 450; Conservative 0; Mismatches 354; Indels 3; Gaps 1;

```

QY 1232 acccaatcgggggtgcgcctgtgtgcaaaagaaacaaatagacag 1291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 68 AGCAAAACATCAGCCCTATTAATTTTACAAACATTTGATTAATCCAGATTTGATG 127
QY 1292 ccgcagtcagaagaaagaaagccagacacatgtgaaggaattagatctcgaaagt 1351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 128 TCAGTCATTTGGAGGTGATGTCACATGAGTGAATGGATTTGATTTGATGATGT 187
QY 1352 cctctctctatcatgtgcgcgcagatgtttcctcctcgtggtatccctcaagat 1411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 188 TCAATTTGTTATCCACAAAGACATCTCATCATTTAAAGAAATTTGACCTTGAAAT 247
QY 1412 tgaagcgaagaaagacagtagcatgtgtgggagcagcgctgtggaaagacacttgt 1471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 248 TAAGAAAGGAGAAATTAATTTGATTGTTGGAGCATCAGAGATTAAGAAATCAACTACTAT 307
QY 1472 tcaactctgcagagacttatagcccgctgcaaggaagtgctgtgtgagtgtgga 1531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 308 TCAATTTATCCAAAGAAATTTATGAAACAAATGGTGAAGAGTAATAGATGGAAGA 367
QY 1532 tgcgaagaagatgaatgacagtggtcgcgtcccaatagaatcgttctcaagagcc 1591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 368 CATAGAGAGAGTTGATATCAATGCTTAAGAAATTAAGAAATGAGTAATTTGACAAAGAAC 427
QY 1592 tgtgtcttcaacgcagcatgtgtggaacatgcctatgtgtgacaacagccgtgtgt 1651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 428 AGTATTTATTTTCAGGACAAATTTGGAGAAATATTCATGCTTGGAGCTTAAGAGGAAGAAC 487
QY 1652 gccac---ttagatgagatcaagaagccgcaaatgcagcaaatatcatcttcttatgta 1708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 488 ACTAAGTAAAGAGAGATGATGTAATTAAGAAATGCAATGCAATGACTTATTTTC 547
QY 1709 aggtctccctgagaatacaacacagaagtgtgagtaagaagagacagcttctgcggg 1768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 548 TAAACTTTCCAAAGAGATTTGACACATATTTGAGAAAGAGAGCATTTATTCAGGAGG 607
QY 1769 ccagaaacaaagactagctatgtcaagggtccttccaaaaacccaaatcttctgt 1828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 608 ACAAAACAAAGAGATTTGCAATTTGACGTGCAATTTGGAAGCCATCATTTCTTCTCT 667
```

```

QY 1829 ggaatgagccacttcagccctcgataatgacagtgaaggtgtcgaagcatgccttga 1888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 TCATGATCATCATATGACGACCTTGATTCACAAAGTGAAGATTTGACAAAGAGCACTTCA 727
QY 1889 taaagccaggaagcgaagacatgcttagtgttactacacagcgtctctgcgaattcga 1948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 CAAGCATCAAAAGAGAAAGAACAAATTTGTAGACATTAACATATCAGATTTCAAAA 787
QY 1949 ccgacattgtagtggtgtgcacatgtgaagaatgaagaagacacacatcaaga 2008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 TCAGATTCAAATATGTGTTATTTATGAGAGAAATAGTAGAACAGAGACATTCAGA 847
QY 2009 gctcctgagaatcgagacatatatt 2035
DB 848 ATTTATTTGATTTGAAAGATTTTAT 874
```

RESULT 14  
 LOCUS A2540627 886 bp DNA GSS 14-NOV-2000  
 DEFINITION ENT018TF Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.  
 ACCESSION A2540627  
 VERSION A2540627.1 GI:11147603  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.  
 ORGANISM Entamoeba histolytica.  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE 1 (bases 1 to 886)  
 Loftus, B., Van Aken, S. and Fraser, C.  
 Determination of clone end sequences from Entamoeba histolytica HMI:IMSS sheared DNA library  
 Unpublished (2000)  
 CONTACT: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: entae@tigr.org

## JOURNAL

## COMMENT

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library  
 Seq primer: M13-Forward  
 Class: shotgun  
 High quality sequence start: 23  
 High quality sequence stop: 856.  
 Location/Qualifiers

## FEATURES

Source

```

1. 886
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db.xref="taxon:5759"
/clone.lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOST1; Site_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaundin and B. Barrell, Oxford University Press, 1999)."
```

## BASE COUNT

364 a 107 c 181 g 234 t

ORIGIN

Query Match 10.7%; Score 221.4; DB 13; Length 886;  
 Best Local Similarity 57.5%; Pred. No. 2.1e-45;  
 Matches 455; Conservative 0; Mismatches 331; Indels 5; Gaps 3;

OY	1267	cggaagaagaacccaattatagacgcgcgcagtcgaagaaggaanaagccaagcaactg	1326
Db	23	TTGATGAAATACCGAATATTTGATTTGATGATCTACAGCTGGATTGATTCACAACTGA	82
OY	1327	aaggaattttagttcttcgaagtcctctctctctctcatctgcccagatgttcca	1386
Db	83	TTGGAAATATTTAAATTTAGAACATTTTCATTTTACATATTCACACAGACCAATTAACAA	142
OY	1387	tcctccgttgcttaccctcagtatgacgcggaagaagacagtagcatgtgtggaagca	1446
Db	143	TTCTGTGTGGACTTGACCTTGAAATTTAAGAAAGSAAACAGTTGCATTGTGAGAGCAT	202
OY	1447	gcgcctgtgggaagaagcaactctgttcaactctgcagagacttctbaccocggtcaag	1506
Db	203	CAGGATGTGGTAAAGTCAACATCAATTTCAATTAATGCCAAAGAGCTCATGTGATCCAGTTGGTG	262
OY	1507	gacaagtcgtttgatgtgtgtgaatgcaagaatgaaatgaaatgacgtgtctctcc	1566
Db	263	GAAAGATACATTTGATGATGGAATATTTAAGAGAGTGAACCTTGAAATGGTTAAGAAATC	322
OY	1567	aaatgcacatgcttccctcacaagccgtgtgctcttcaactcgaacatgtcgaagaacatcg	1626
Db	323	AAATAGGATTTAGTTGGACAAAGAACCTATTGTTTGATGTACTATTAGGAATAATTA	382
OY	1627	cctatgtgtacacacagccgtgtgtgcc---ttagaatgacaaagaagccgcaatg	1683
Db	383	TGCTTTGGAGCTTAAAGATGAGGAAACACCAACTGMAAGAGATCATTTGAATGTGCTTAAA	442
OY	1684	cagcaaatcatctcttttatitgaagtctccctgtgaataataacacacaaagttggac	1743
Db	443	TGGCAATGCACTGATGAAATTTATTTCTCATCTCCAGAGAGATATGATACAAATGTAAGAG	502
OY	1744	tgaaagagcagcagcttctgtgcgcagagaacaaagaactgcattcgaaggcccttc	1803
Db	503	AAAGAGGAGCTGCAATTTACAGAGAGCAAAACAAANAATTTGCTATTGCACGTGCAATTA	562
OY	1804	tccaaaaaccaaaatttatltgttgaagaagccacttcagccctcgatataatgacagt	1863
Db	563	TTAGAAACCTTACATTTATTTATTTACTGTGAAGCTACATCAGCATGTAATACAAAGTG	622
OY	1864	agaaggtgtgtcagcatgccccttgataagcccggaagcgaaggaacatgctgtgtgca	1923
Db	623	AAAGATTGTACAACCAAGCACTTGAAAAACCTAGTCAAGGAAACAAACAAATTTGTGAG	682
OY	1924	ctcagagcctctcgcgaattcagaagcagatgttagtctg--gttctgcacaatgaag	1982
Db	683	CACATAGATTAAACAACGTTTAGAAATGCCAATAGAAATTTGTGTATTTCCATCAAGAGAA	742
OY	1983	ataaagaacaagaactcatcaagaagctccttgaagaatcgagacatalatttaagta	2042
Db	743	ATTATTGTAACAAGAACATCAATCAAGATTATATG--GATTTGAAGCACATTTATTTGATTA	801
OY	2043	gtgaatgcaca	2053
Db	802	GTTAAAGACA	812
RESULT	15		
LOCUS	AZ541090	897 bp	DNA
DEFINITION	ENTDS677R Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.	GSS	14-NOV-2000
ACCESSION	AZ541090		
VERSION	AZ541090.1	GI:11148493	
KEYWORDS	GSS.		
SOURCE	Entamoeba histolytica.		
ORGANISM	Entamoeba histolytica		
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.		
AUTHORS	1 (bases 1 to 897)		
TITLE	Loftus,B., Van Aken,S. and Fraser,C.		
	Determination of clone end sequences from Entamoeba histolytica		
	HHI:IMS shared DNA library		
JOURNAL	unpublished (2000)		

COMMENT	CONTACT: BRENDAN J LOFTUS
	Department of Eukaryotic Genomics
	The Institute for Genomic Research
	9712 Medical Center Dr., Rockville, MD 20850, USA
	Tel: 301 838 0200
	Fax: 301 838 0208
	Email: entae@tigr.org
	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
	DNA library
	Seq primer: M13-Reverse
	Class: shotgun
	High quality sequence start: 81
	High quality sequence stop: 890.
	Location/Qualifiers
FEATURES	1..897
source	/organism="Entamoeba histolytica"
	/strain="HMI:IMSS"
	/db_xref="taxon:5759"
	/clone_lib="Entamoeba histolytica Sheared DNA"
	/note="vector: pHOSt1, Site 1: Bst I; Constructed at The
	Institute for Genomic Research (TIGR), Rockville, MD.
	Genomic DNA isolated from broth cultures of E. histolytica
	using a method described by Clark and Diamond (Clark,
	C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
	method for isolate identification. Exp. Parasitol.
	77:450.). The DNA was mechanically sheared to give a
	tight size distribution (~2 kb). The v + i method used for
	the library construction is described in detail in Smith
	H.O. and Venter, J.C. (Making small insert libraries for
	whole genome shotgun sequencing projects. In Genome
	Sequencing: A Practical Approach, eds. M. Vaundin and B.
	Barell, Oxford University Press, 1999)."
BASE COUNT	335 a 110 c 177 g 275 t
ORIGIN	
Query Match	10.7%; Score 220.4; DB 13; Length 897;
Best Local Similarity	57.2%; Pred. No. 3.8e-45;
Matches 420; Conservative 0; Mismatches 311; Indels 3; Gaps 1,	
QY 1325 tgaagggaattagaggtttcgagaagctctcttctcattccatgctgcccaatgcttt 1384	
DB 93 TAAAGTGAATTTGAATTTAAAGACATTTGTTTCAGATATCCAAACAGACGACAAATTC 152	
QY 1385 catctccgtggtctatccctcagatcatgagaggaagacagatgagcttggggg 1444	
DB 153 TGCTTTGAAGGTAATTTCTTCGAAAGTGAACAGAGAAAACGTTGCTATGTAGAGACC 212	
QY 1445 cagcgctgtggaagaagcaactctgctcaactctgcagagagacttiagaccgcgtgca 1504	
DB 213 ATCAGAGATGTGTAATCAACATCAGTTCAAGTTGAAGATTTTATGATCAACAACA 272	
QY 1505 aggcacagtgctgtttagatggtgagatgcaagaattgaatgacagtgctccgttc 1564	
DB 273 TGGACATGTATTTAGAGATGACATATATTCAAAGATTTGAATATTCATTTCTTAAAG 332	
QY 1565 ccaaatagacatcgttctcctcaagagccttgctcttcaactcagcatgctgcgagaaat 1624	
DB 333 TCAAAATTTGGAATGTGATGACCAAGAACCACTATTATTTGCTGAAAGTGTTATGATATAT 392	
QY 1625 cgactatggtg--acaaacagcgtgtgtggtccattagatgagatcaagaagccgcaaa 1681	
DB 393 TGAAGAGAGATACCTAAAGAGATTGAAGTATGATGATCAACAATTTATGCTGCTGCTAA 452	
QY 1682 tgcagcaaatatccattcttattgaaggtcttccctgagaataacacacaagtgttg 1741	
DB 453 AATGCAATATGCATGACTTTATTTACCAATGCGAGAGATATTAACACAAATGCTAG 512	
QY 1742 actggaagagacagctcttgcggcgccaagaacaaagactgactatgcaaggctct 1801	
DB 513 TGATGAGAGTGCACAAATTTTCAGGAGGACAAATAAAGCAATTCGCTATTGTCACGTGCATT 572	
QY 1802 tctccaaaacccaataatttattgttggatgagggccacttcaagccttcgataatgacag 1861	

```
Db 573 GATTAGAAATCCAAAAGTGTATTACTCGATGAGCTACATCAGCACTTGATTGAGAAAG 632
Qy 1862 tgaagaaggttggttcagcatgccccttgataaagccaggaaggaagacatgcccagtggt 1921
Db 633 TGAAGAAGATTGTACAGAGATGCACTTGACAAAGCAGCAAAAGGAAGAACAAATTTGTAAT 692
Qy 1922 cactcaaggctctctgcaattcagaacgcagattgtagtgggtctctgcacaatgga 1981
Db 693 TGCACATAGATTATCAACTATTCAAAATGCAGATCAAAATATGTGTATTATGAGAGGAAG 752
Qy 1982 gataaaggaaagaaactcacaagaagctcccgagaatcgaacatataatatttaagtc 2041
Db 753 AATTGCAGAAAGAGACACATCAAGAGTTATTAGATTGGAAGGATTTATTATTACACT 812
Qy 2042 agtgaatgcacagt 2055
Db 813 TCCTATGCACCAAT 826
```

Search completed: April 22, 2002, 18:31:37  
Job time: 12190 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 15:10:17 ; Search time 293.81 Seconds

(without alignments)  
1592.538 Million cell updates/sec

Title: US-09-873-409-9

Perfect score: 2066  
Sequence: 1 cgaagcaaggtcgactaca.....atgcacagtcagtcagtgta 2066

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents: NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCITUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/Backfile1.seq: \*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	712.4	34.5	2726	1 US-08-461-823-1	Sequence 1, Appli
2	712.4	34.5	4666	1 US-08-181-471-2	Sequence 2, Appli
3	712.4	34.5	4669	6 5206352-3	Patent No. 5206352
4	712.4	34.5	6505	2 US-08-793-610-5	Sequence 5, Appli
5	712.4	34.5	9318	2 US-08-793-610-6	Sequence 6, Appli
6	710.8	34.4	4264	2 US-08-784-649A-1	Sequence 1, Appli
7	710.8	34.4	4264	2 US-08-784-649A-5	Sequence 5, Appli
8	710.8	34.4	4669	2 US-08-583-276-18	Sequence 18, Appli
9	704.4	34.1	4669	2 US-08-752-447-1	Sequence 1, Appli
10	645	31.2	4233	3 US-09-120-513-1	Sequence 1, Appli
11	645	31.2	4233	4 US-09-450-105-1	Sequence 1, Appli
12	285.8	13.8	4047	2 US-08-612-734B-1	Sequence 1, Appli
13	283.8	13.7	4002	2 US-08-996-545-1	Sequence 1, Appli
14	283.8	13.7	4002	2 US-08-996-545-3	Sequence 3, Appli
15	283.8	13.7	4002	4 US-09-328-320-1	Sequence 3, Appli
16	283.8	13.7	4002	4 US-09-328-320-3	Sequence 3, Appli
17	238.8	11.6	3924	1 US-08-395-246C-1	Sequence 1, Appli
18	222	10.7	4224	1 US-08-612-521-1	Sequence 1, Appli
19	213.8	10.3	4800	2 US-08-612-734B-3	Sequence 3, Appli
20	145.2	7.0	2407	2 US-08-895-522-2	Sequence 2, Appli
21	145.2	7.0	2407	3 US-09-195-391-2	Sequence 2, Appli
22	145	7.0	5120	3 US-08-772-270A-6	Sequence 6, Appli
23	145	7.0	8370	2 US-08-488-706-1	Sequence 1, Appli
24	144.8	7.0	7721	3 US-08-772-270A-14	Sequence 14, Appli
25	141	6.8	3924	2 US-08-996-644-3	Sequence 3, Appli
26	141	6.8	3924	3 US-09-352-552-3	Sequence 3, Appli
27	141	6.8	3927	2 US-08-996-644-1	Sequence 1, Appli

28	141	6.8	3927	3 US-09-352-552-1	Sequence 1, Appli
29	133.4	6.5	1201	3 US-08-961-083-189	Sequence 189, App
30	133	6.4	2244	4 US-09-061-764A-18	Sequence 18, Appli
31	132.6	6.4	2061	4 US-09-061-764A-17	Sequence 17, Appli
32	130.4	6.3	4403765	4 US-09-103-840A-2	Sequence 2, Appli
33	128.2	6.2	3909	1 US-08-232-537-1	Sequence 1, Appli
34	121.6	5.9	2376	1 US-08-394-880B-1	Sequence 1, Appli
35	115.2	5.6	6143	1 US-08-612-521-3	Sequence 3, Appli
36	114.8	5.6	1959	4 US-09-061-764A-4	Sequence 4, Appli
37	110	5.3	28804	2 US-08-592-874-1	Sequence 1, Appli
38	110	5.3	28804	3 US-09-096-942-2	Sequence 2, Appli
39	110	5.3	28804	3 US-09-096-867-2	Sequence 2, Appli
40	102.6	5.0	4403765	4 US-09-103-840A-2	Sequence 1, Appli
41	102.6	5.0	4411529	4 US-09-103-840A-1	Sequence 1, Appli
42	80.6	3.9	5889	1 US-08-463-092B-5	Sequence 5, Appli
43	80.6	3.9	5889	2 US-08-462-109A-5	Sequence 5, Appli
44	80.6	3.9	5889	2 US-08-460-907B-5	Sequence 5, Appli
45	80.6	3.9	5889	3 US-08-463-179A-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-08-461-823-1  
; Sequence 1, Application US/08461823  
; Patent No. 5353840  
; GENERAL INFORMATION:  
; APPLICANT: Bhattachar, Satish K.  
; APPLICANT: George Jr., Albert L.  
; APPLICANT: Nazarenko, Irina  
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: OncorPharm, Inc.  
; STREET: 200 Perry Parkway  
; CITY: Gaithersburg  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20877  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08461,823  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/168,621  
; FILING DATE: 16-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/010,433  
; FILING DATE: 27-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Karta, Glenn E.  
; REGISTRATION NUMBER: 30,649  
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301 527-2058  
; TELEFAX: 301 208-6997  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2726 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; US-08-461-823-1











OY	262	acttccatcgaagcgctcgtgagaaatccaccacaatcaaaagatagctcttcgaagct	321
Db	3843	AAGCCCAAGACAGAAAGCTTAGTACCAAAAGAGGCTTGATGAAAGTATACCTCCAGTT	3902
OY	322	ctccattaaaaatttlaaagtlaaacaagccgaaatgsgctcttggcttcggggagcat	381
Db	3903	CCTTTGGAGAGATTATGAACTTAATTTAACTGAATAGGCTTATTTGGTGTGGTAT	3962
OY	382	tgccttcgtcttaaatgaaactgcttcatccagtaatttccatcactcttgcanaaalia	441
Db	3963	TTTTGGCAATTAAATGAGGCGCTCAACACGACATTTGCATTAATTTTCAAGATT	4022
OY	442	taaccagcttggaaataaagataaaacacatitiaag---cagtgacgaataattat	498
Db	4023	TAGGGGTTTTTACAGAAATTGATGATCCGTGAACCAACGACGAATATGTAACCTTGTTT	4082
OY	499	ccaagatattcgatcaatttgggtgtaatttgccttcttgcagttaatttcagaggaattat	558
Db	4083	CACATTTGTTTCAGCCCTTGGAATTAATTTCTTTTATACATTTTCTCTCAGGGTTTCA	4142
OY	559	ttacgscagagcaggggaaatttlaacgatagaattaaagacatttggccttcaagaca	618
Db	4143	CATTGGCAAAAGCTGAGAGATCCTCACCAAGGGCTCGATACATAGGTTTTCCGATCA	4202
OY	619	tgtatatacggatattgctcttggtttgaagaaaggaacacgcacagaggtcttgcaca	678
Db	4203	TGCTCAGACAGGATGTAGTTGGTTTGATGACCTTAATAAAACACCACTGGAGCATGACT	4262
OY	679	caatattagccaatagatatagcaacaattcaagagcaacaggttccagagattgsgct	738
Db	4263	CCAGGCTCGCAATGATGCTGCTCAAGTTTAAAGGGCTTTAGGTTCCAGGCTTGCTGTA	4322
OY	739	taacacaaaatgacactaaacatgysacttcagtaattcacttcccttatalatagatbg	798
Db	4323	TTACCCAGAAATATAGCAAACTTGGGACAGAAATAATTATTCCTTCACTATGGTTGC	4382
OY	799	agatgacattccctgattctcgtatgattgctccagacttgcgltgacaggaatgatitgaa	858
Db	4383	AACTAACACTGTTTACTCTGTGCAATTTGTACCCATCTTTCATTAACAGAGTTGTTGAA	4442
OY	859	ccgcagcaatgacatgagatttgcaccaaaaagataaagaagaaacttaagacttgcitgaaga	918
Db	4443	TGAAATGTGTCTCGACACAGCACTGAAGATTAAGAAATACTACAGAGTCTGGGAAGA	4502
OY	919	taagcaactgaagcttggagaaataacgttaactaagtgtcatttlaacaaaggganaaacct	978
Db	4503	TCCGTACTGAAAGCAATATGAAAACTTCCGAACCTGTTCTTTGACTCAGAGCGAGAGT	4562
OY	979	tcgcgcaaatgataagaagatagcttcagatcacaacagaatatccctcgaaagaagac	1038
Db	4563	TTTGAACATATGTATGCTCAAGATGTTTGCAGGTACCATTAAGAAATCTTTGAGGAAGAC	4622
OY	1039	agattattggaagctgttgaatcatcaagcatcgcatgccccttatalatttgcatagcag	1098
Db	4623	ACATCTTTGGAATTACATTTTCTTCAACAGCAATGATGATTTTTCCTTTCGTGAT	4662
OY	1099	gatttcgatttggagccatttlaatcaagacttgcagaaatgaaaccccgagagggcagtltca	1158
Db	4683	GTTTCCGGTTTGGAGCCTACTTGTGTGGACATTAACATCATGAGCTTGGAGATGTTGCT	4742
OY	1159	tagtttttaactgtcaattgcatatgagactatgcatcgggaaaaacgctgttgyctc	1218
Db	4743	TAGATTTTTCAGCTGTGTGTTTGGTGCCATGGCCCTGGGGCAAGTCATTTATTTGCTC	4802
OY	1219	ctgaaatccaaagccaatcgggggctgysgactgttgccttgttggaaagaaac	1278
Db	4803	CTGACTATGCCAAAGCCAAAATATCACAGACCCACATATCATATGATCATTTGAAAAACC	4862
OY	1279	caaatatagacacgcgcagtcaagaagggagaaagccaaacacatgtgaagaaatttag	1338
Db	4863	CTTTGATTGACAGCTACAGCAGGAAGGCCCTATATGCCGAACACATTTGGAAGGAATGTCA	4922
OY	1339	agtttcggagaagctcttcttctcatalcatatgltccacagaatgtttcaatccctcgltgc	1398

Dd	4923	CATTGGTGAAGTGTATTTCATACTATCCACCCGACCGGACATCCCAATGCTTCAAGSAC	4982
Qy	1399	tatccctcaatgatctgagcaggaagaacagatgacatttgtyggagcagcgctgtgga	1458
Dd	4983	TGAGCGTGGAGGTGAAGAAAGGCGCAGCGCTGGCTGTGTGGGACGAGTGGCTGTGGGA	5042
Qy	1459	aagacactctgtctcaactctctgcagagacttatbgaaccggtgcaagagacaagtgtcgt	1518
Dd	5043	AGACACAGTGGTCCACTCTCTGGACCGGTTCACGACCCCTTGGCAGGAAAGTGTGC	5102
Qy	1519	ttagtggtgtagatgcacaaagaattgaatgacagtggtctccgtccccaatagcaatcg	1578
Dd	5103	TTGATGGCAAGAAATTAACCGACTGAATGTTCAATGAGCGCTCCGACACACTGGGCATCG	5182
Qy	1579	tctctcaagagcctgtgctctcttcaactgcagcaatgctgtgagacacatgcctatggtgaca	1638
Dd	5183	TGTGCCAGAGACCCACTCTGTGTTGACTGCAGCATTCCTGAGAACTTGCCTATGAGACA	5222
Qy	1639	acagccgctgtggtgcatctatgatatgaataagaacgcgcaaatgcagaaatattccatt	1698
Dd	5223	ACACCGGGGTGGTCTCACAGGAAGAGATCGTGAAGGCCACAAAGAGGCCCAACATACATG	5282
Qy	1699	ctttattgaaggtctccctcgagaaataacacacacaagtgtgactggaagaagacagc	1758
Dd	5283	CCTTATATGAGTACTCTCCCTTAATTAATATBGCCTTAAGTATGAGACAAAGAACTCAGC	5342
Qy	1759	tcttcgcgcgcacaaacaaagaactgactatltgcaaggctctctccaaaaccacaaa	1818
Dd	5343	TCTCTGGTGGGCNAACAAACAGCATTCGCACTAGCTGTGTCCTGTTTGAGACGCTCATTA	5402
Qy	1819	tttattggttgatgagcgcaacttcagccctcgataatgacagtgagaagtggtgtcagc	1878
Dd	5403	TTTTGCTTTTGGATGAAAGCCACGCTGACCTGTGATACAAAGATGMAAAGTTGTCCAAAG	5462
Qy	1879	atgccccttgataaagccaggaacggtgaagacatgcttagtggtgaactcaagcgtctgtg	1938
Dd	5463	AAGCCCTGGACAAAGCCAGAAAGGCCGCGACCTGCAATGTGATGCTACACGCGCTGTCA	5522
Qy	1939	caattcgaacgcagatttatagtggttctgcacatggaagaataagagacaaagaa	1998
Dd	5523	CCATCCAGAAATGCGAGCTTAAATAGTGTGTTCAGAAATGGCAGATGCAAGAGCATGCGCA	5582
Qy	1999	ctcatcaagagctccctgtagaatacgcagacatatatttaagttagtgaatgcagcttag	2058
Dd	5583	CGCATCAGCACTCTCTGCGACAGAAAGCATCTATTTTTCAATGGTCAGTGTCCAGGCTG	5642
RESULT 5			
US-08-793-610-6			
; Sequence 6, Application US/08793610			
; Patent No. 5858744			
; GENERAL INFORMATION:			
; APPLICANT: BAUM, Christopher			
; APPLICANT: STOCKING-HARRERS, Carol			
; APPLICANT: OSTERING, Wolfertm			
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF			
; NUMBER OF SEQUENCES: 6			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP			
; STREET: 655 Fifteenth Street N.W. Suite 330			
; CITY: Washington			
; STATE: D. C.			
; COUNTRY: U. S. A.			
; ZIP: 20005-5701			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION NUMBER: US/08/793,610			





D	2945	ACATCTTGGAAATTACATTTTCTTCCATTACCCAGGACATGATGATTTTTCCTATGCTGGAT	3004
Q	1099	ggttcgatttggagccatttaataatcaagcttgcagaaatcccccaaggaggtcatttca	1158
D	3005	GTTTCCCGTTTGGAGCCTTACTTGGTGGCACATTAACATCATGAGCTTTGAGATGTTCTGT	3064
Q	1159	taattttacttgcgaattgcataatgagagctatgagccatgcgaaaaacgcgttttgctc	1218
D	3065	TAGTATTTTAACTGCTGTGTTCTTTGGTGCCATGGCCGTGGGCAACATGATTCATTTGCTC	3124
Q	1219	ctgaatatccaaagccaaatcgggggctgcgcattcgtttgccttgcgttgcgaagaagac	1278
D	3125	CTGACTATGCCAAAGCCAAATAATATCAGACAGCCCACTATCATGATCAATTGAAAAAACCC	3184
Q	1279	caaatatagacagccgcagtcctcaagaagggaaaaaagccagacacatgttgaaggaattag	1338
D	3185	CTTTGATTTGACACTACAGACAGGAAGGCCCTAATGCCGAACCAATTGGAAGGAATGTCA	3244
Q	1339	agtttcgagaagctcctcttctatccaatctgcgcccaatgttttcaatccctcgttgc	1398
D	3245	CATTTGGTGAAGTTGTTATTCATATCCACCCGACCCGACATCCCATGCTTCAAGGAC	3304
Q	1399	tatccctcaglatitgcgcgaagaagaacagtagcatttgcgttgcgcagcgcgttgcga	1458
D	3305	TGACCCCTGGAGGTGAAGAAGGCCCAACCCCTGGCTCTGTTGGGCGACAGTGGCTTGCGGA	3364
Q	1459	aaagcactctcttcaactcttcacagagcttatagcccgctgcagaagaacagtcgt	1518
D	3365	AGACACANGTGGTCCAGCTCTCTGGACGGTTCTACACCCCTTGGCAGGGAAGTGGTGC	3424
Q	1519	ttgatgtgcgtgcgtgcgaagaagaattgaatgtacagtgctccgttcccaatagaatctg	1578
D	3425	TGTATGGCAAAAGAAATAAACCGACTGAATGTTTCAGTGGCTCCGACACACCTGGGCAATGC	3484
Q	1579	ttctcgaagggctcgtgctcttcaactgcagcattgcttgcagaactctgcctatgttgca	1638
D	3485	TGTCGCCAGGAGCCCATCTCTGTTTGACTGGACGATTTGCTGTAGAACATTTGCTATGGAGCA	3544
Q	1639	acagccatgtgcgtgcacattgaatgcagatcaagaagccgcaaatgcacaaatatacat	1698
D	3545	ACAGCCGGGTGGTGTCCACAGAGAAATTTGTGAGGCGACCAAGAGAGGCCAATACATG	3604
Q	1699	ctttatitgaagtlctccctgagaaatatacacacacaagtlygacttgaagagcacagc	1758
D	3605	CCTTCACTGAGTCACTGCCCTTAATAATATATAGCACTAAATAGAACACAAAGAACTCAGC	3664
Q	1759	tttttcgcgcgcggaagaacaaagctcgtcatitgcgaaggtcttctccaanaaccacaa	1818
D	3665	TCTCTGGTGGCCAGAAACAACGCGATTGGCCATATGCTGTGCCCTTGTAGACAGCCTCTA	3724
Q	1819	ttttatgttgcgtgcgtgcacacttcagccctcgataatcacagttagaagtggttcgc	1878
D	3725	TTTTGCTTTTGGATGAAGCACGCTCAGCTCTGGATATCAAAAGTGAAGAAAGTGTTCAG	3784
Q	1879	atgcgccttgaataagccagaagcgggaagacatgctagttgtcactcaacagctctctg	1938
D	3785	AAGCCCTGACAAAGCCAGAGAAAGGCCGACCTGCACTTGTGATTCCTCAACCCCTGTCCA	3844
Q	1939	caatcgaacgcagaaatttatgtgtgtctgcacatgcgaagaataaaggaacgaagaa	1998
D	3845	CCATCCAGAAATGCAAGCTTAATATGTGGTGTTCAGAAITGGCAGATGCAAGACATGGCA	3904
Q	1999	ctcatcaagagctccctgcagaataatcgagacatatatttaagtttagtgcagtcagtcag	2058
D	3905	CGATTCAGCAGCTCTGGCCACAAAGGCGATCTATTTTCAATGGCTCATGTGCCGCTG	3964

RESULT 7  
US-08-784-649A-5  
; Sequence 5, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:

1 APPLICANT: Sikic, Branimir I  
 2 APPLICANT: Chen, Gang  
 3 TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
 4 TITLE OF INVENTION: CYCLOSPORIN MODULATION  
 5 NUMBER OF SEQUENCES: 5  
 6 CORRESPONDENCE ADDRESS:  
 7 ADDRESSEE: Fish & Richardson  
 8 STREET: 2200 Sand Hill Road  
 9 CITY: Menlo Park  
 10 STATE: CA  
 11 COUNTRY: USA  
 12 ZIP: 94025  
 13 COMPUTER READABLE FORM:  
 14 MEDIUM TYPE: Floppy disk  
 15 COMPUTER: IBM PC compatible  
 16 OPERATING SYSTEM: PC-DOS/MS-DOS  
 17 SOFTWARE: Patentln Release #1.0, Version #1.25  
 18 CURRENT APPLICATION DATA:  
 19 APPLICATION NUMBER: US/08/784,649A  
 20 FILING DATE:  
 21 CLASSIFICATION: 435  
 22 ATTORNEY/AGENT INFORMATION:  
 23 NAME: Sherwood, Pamela J  
 24 REGISTRATION NUMBER: Reg.No. 5830697 36,677  
 25 REFERENCE/DOCKET NUMBER: 06037/007001  
 26 TELECOMMUNICATION INFORMATION:  
 27 TELEPHONE: 415-322-5070  
 28 TELEFAX: 415-854-0875  
 29 INFORMATION FOR SEQ ID NO: 5:  
 30 SEQUENCE CHARACTERISTICS:  
 31 LENGTH: 4264 base pairs  
 32 TYPE: nucleic acid  
 33 STRANDEDNESS: single  
 34 TOPOLOGY: linear  
 35 MOLECULE TYPE: CDNA  
 36 US-08-784-649A-5

[illegible]





LENGTH: 4669 bases  
TYPE: nucleic acid  
STRANDEDNESS: singular  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: Genomic DNA  
US-08-583-276-18

Query Match 34.4%; Score 710.8; DB 2; Length 4669;  
Best Local Similarity 60.0%; Pred. No. 1.9e-204;  
Matches 1261; Conservative 0; Mismatches 797; Indels 42; Gaps 3;

Oy 1 cgaagaagaatgcgagacacatcgtgtaagcaaccgactctactactatctgaatgcag 60  
Db 2151 CCAGAAAAGTGCGACACCATGTGATAGCTCATCGTTTGTCACAGTTCCTAATGCTG 2210  
Oy 61 attgatgtacccctaaagatgtaatgctgcgagaagaagagacacatgctgaactaa 120  
Db 2211 ACGTCATGCGTGGTTTGCATGATGAGTCATTTGTGAGAGAAAGAAATCATGTGAACCTCA 2270  
Oy 121 tggcaaaacgaggtctatattactactgtgtaatgtaacaggaatataaaagctgaty 180  
Db 2271 TGAAGAGAAAAGGCATTACTTCAAACTGTGCACATGTCACAGACACAGAGAAATGAAGTTG 2330  
Oy 181 aa-----cagatgagtcgaatgacatatctactcagaagaagaaccaactca 227  
Db 2331 AATTGAAAATGACGCTCATGTGAATCCAAAAGTGAATTTGATGCTTGGAAATGCTTCA 2390  
Oy 228 ctccctgcacctcgtgaagacatcaagtcagacttcattgacaagctgaggaatcc 287  
Db 2391 ATGATTCAAGATCAGCTCTAATAAGAAAAGATCAACTGCTGAGATGTCCTGGATCAG 2450  
Oy 288 aacc-----aattcaagaagaatgaagcttccctgaagct 321  
Db 2451 AACCCCAAGACAGAAAGCTTAGTACCAAAAGGCTCTGAGTGAAGTATACCTCCAGTTT 2510  
Oy 322 ctctatataaatttaagtaaaacagcctgaatgaccttctgtgtctcgggagact 381  
Db 2511 CCTTTGAGAGGATTATGAAAGCTTAATTTTAACTGGAATGGCTTATTTTGTGTTGGTAT 2570  
Oy 382 tggcttcgtctaaatggaactgtctccagatatttccatctctcttgaataaata 441  
Db 2571 TTTGTGCATTAATAATGAGGCGCTGCAACGACATTTCCATTAATTTTCAAAAGATTA 2630  
Oy 442 taaccatgtctggaaataatgataaaaccacataaag---caltgacagaatattat 498  
Db 2631 TAGGGGTTTTTACAGAAATTTGATGATCCGAAACAAACGACAGAAATACTACTGTGTTT 2690  
Oy 499 ccatgatatctgcatcttgggtgtatttggctgtcagatattcattcagcaggatatt 558  
Db 2691 CACTATTGTTCTAGCCCTTGGAATTAATTTCTTTATTAATCAATTTTCCCTGAGGTTTCA 2750  
Oy 559 ttacgcagagacgagggaaattttaacgatgataaagacactgagcctcacaagcca 618  
Db 2751 CATTTGGCAAGAGCTGAGAGATCCTACCAAGGCGCTCGATACATATGTTTCCCATCCA 2810  
Oy 619 tgtatcatcaggaatatgctcgtgttctgtaagaagaagaacagcacaagagcttgaca 678  
Db 2811 TGCCTAGACAGAGATGTGATTGGTTTGATGACCTTAATAAACACCACTGAGCATTTGACTA 2870  
Oy 679 caatatgacatagatatagacaatccaagagcagaagcagctccagatctgcgtct 738  
Db 2871 CCAGGCTCGCCATATGATCTGCTCAAGTTAAAGGGGCTATAGGTTCCACAGGTTGCTGTA 2930  
Oy 739 taacacaatatgcaactaacatgagacttcagttatcatctccttatataatgagtg 798  
Db 2931 TTACCCAGAAATATAGCAATTTCTTTGGACAGAAATATATATATATCTTCACTATGTTGCC 2990  
Oy 799 agatgacatctcgtatctgagtatgtctccagtaactgtgcgtgacgaagaatgatgaa 858  
Db 2991 AACTAACACTGTACTTACCAATTTGTAACCATATGCAATTCATAGCAGAGATTGTTGAAA 3050

Oy 859 ccgcagcaatgacgtgatttgcacaagaagataagcaagaacttaagcatgctggaaga 918  
Db 3051 TGAATAATGTTGCTGTGACAAAGCATGAAAGATTAAGAAAGAACTAGAGAGTGTGGAA 3110  
Oy 919 tagcaactgaagccttctggaatatacgtactatagtgctcatatacaagaagaagcct 978  
Db 3111 TCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTTCTTTGACTACAGAGACAGAGT 3170  
Oy 979 tcgagcaaatgtaataagagatgcttcagactccaacagaataaacctcagaagaagcc 1038  
Db 3171 TTGAACATATGTATGCTGAGTGTGACAGTACCATACAGAAACCTTTGAGGAAGAGC 3230  
Oy 1039 agattatggaagcctgtatgcatcagcactgagccttataatttgcctatgacagcag 1098  
Db 3231 ACATCTTTGGAATTTACATTTTCTTCCACCAAGCAATGATGATTTTCTTATGCTGGAT 3290  
Oy 1099 ggttcgaatttggagccttataatcaagcttgaagaatgaccccaagaggaatgtca 1158  
Db 3291 GTTTCCGTTTGGAGGCTACTTGTGACATTAATTAATGAGCTTTGAGATGTTCTGT 3350  
Oy 1159 tagctttacgtcaatgcataatgagcctatgagcctcgaagaagaagcctcgtttgccc 1218  
Db 3351 TAGTATTTTACACTGTTGTTGTTGTCATGGCCCTGGGGCAGAGTTCATTTTGTCTC 3410  
Oy 1219 ctgaatattccaagaacaaatcgagggtcgcatctgttgccttctgttggaaagaagaac 1278  
Db 3411 CTGACTATGCCAAGCCAAATATACAGCCACCCATCATGATCATTTGAAAAACCC 3470  
Oy 1279 caataatagacagccgacatgcaagaagaagaagaacacatgtlgaaggaaattag 1338  
Db 3471 CTTTGTATGACAGCTACAGCAGAGGAGGCTTAATGCCGAACATTTGGAAGAAATGTCA 3530  
Oy 1339 agttcgaagaagctctcttctcatalccatgtccgagaatgtttcatctcgtgct 1398  
Db 3531 CATTTGGTGAAGTTGTTTCACTATCCACCCGAGCAGATCCCATCCATTTTCAAGGAGC 3590  
Oy 1399 tatccctcagtatgagcagaagaagaacagtagcatttggggaagcagcgtctggga 1458  
Db 3591 TGAAGCTTGAGGTGAAGAAAGGCGACAGCCTGACTGTGGGGACAGAGTGTGGGA 3650  
Oy 1459 aaagacactctcttcaactctgcagaagacttaatgaaccgtgcagaaggaacatgctgt 1518  
Db 3651 AGAGCAACATGTGTCAGCTCTGTGAGAGCGTTTACAGACCCCTTGGAGGAAAGTCTCG 3710  
Oy 1519 ctgattgtgtgagatcgaagaagaatgaaatgtaagtgctcgtctcccaaatagaatcg 1578  
Db 3711 TTGATGGAAGAAATTAAGCGAGCTGAATGTTCAAGTGGCTCGAGCAGACCTGGCATCG 3770  
Oy 1579 ttccctcaagagcctgtgctctcaactgagcactgtctgagaacatcgactatggtgaca 1638  
Db 3771 TGTCCAGAGCCCATCTGTTTGTGACTGACAGCTTGTCTGGAACATTTGCTTATGACACA 3830  
Oy 1639 acagccgtgtggtgcatatagatgaagaagaacgcgaatgcaaatatcatt 1698  
Db 3831 ACAGCGGGTGGTGTGACAGAAAGAGATCGTGAGGGCAGCAAAAGGACCAATCATCAG 3890  
Oy 1699 ctttatgaaaggtctcctctgagaataacacaacaagaatgtgactgaaagaagcagc 1758  
Db 3891 CCTTATTCAGACACACCTAATTAATATAGCACTTAATGAGAGCAAAAGGAATCACTCA 3950  
Oy 1759 ttctcgtgcgagaaagaagaactagctatgcaagggcctctcccaaaaacccaataa 1818  
Db 3951 TCTCTGTGTGCGAAGAACACAGCATTTGCCATAGCTGTGTGCCCTTGTGTAGACAGCTCAT 4010  
Oy 1819 tttaattgttgaatgagacacttcagcctcgaataatgacagtgaagaagtggttcagc 1878  
Db 4011 TTTTGTCTTTGGATGAAGCCAGCTGAGCTGTGATACAGAAAGTGAAGAGTTGTCCAAG 4070  
Oy 1879 atgccttgataaagcgaagaaggaagacatgctagtgtaactcaagcagctctcg 1938  
Db 4071 AAGCCTTGACAAACCCCAAGAGAGGCGGACCTGCTATTTGTGATTTGCTACCGGCTTCCA 4130  
Oy 1939 caattcagaacgagatttgatgagtggtctgacaaatggaagaagaagaagaagaa 1998



Db 4131 CCATCCAGATGACAGCTTAATAGTGTCTTCAGATGCGACAGTCAAGCAGCATGGCA 4190  
Qy 1999 ctcataaagctccctgtagaatacgcagacatatatttaagttaagtcacagctcag 2058  
Db 4191 CGCATCAGCAGCTGCTGCGACACAGAAAGCATCATTTTTCATGATGTCAGTCCACCTG 4250

RESULT 9  
US-08-752-447-1

; Sequence 1, Application US/08752447  
; Patent No. 5984088  
; GENERAL INFORMATION:  
; APPLICANT: Mechelner, Eugene  
; APPLICANT: Roninson, Igor B  
; TITLE OF INVENTION: Methods and Reagents for Preparing and  
; TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.  
; STREET: 300 South Wacker Drive, Seventh Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,447  
; FILING DATE: 15-NOV-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NO. 5994088aan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 95,1121  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-9808  
; INFORMATION FOR SEQ. ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4669 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: 1..424  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 425..4264  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: 4265..4669  
; US-08-752-447-1

Query Match 34.18; Score 704.4; DB 2; Length 4669;  
Best Local Similarity 59.9%; Pred. No. 1.6e-202;  
Matches 1257; Conservative 0; Mismatches 801; Indels 42; Gaps 3;

Qy 1 cgagcaaggtcgcagactcaatcgtagagcaacacgacttctactcttgaagtgag 60  
Db 2151 CCAGAAAGGTCGACCCATTTGATAGCTCATGTTTGTACACATTCGTAATGCTG 2210  
Qy 61 atttgaattgacacctaaagatgaatgctgtagaagaagacacatgctgaactaa 120  
Db 2211 ACGTATGCTGCTGTTTCATGATGAGATCATTTTGTGAGAAAGAAATCATGATGACCA 2270  
Qy 121 tggcaaaacagagtcataattatctacttgtagtgcacagagataataaaagctgatg 180

Db 2271 TGAAGAGAAAGCATTTTCTTCAACTTGTCACAAATGACAGACAGCAAGAAATGAAGTTG 2330  
Qy 181 aa-----cagatgagatcaatgacatattctactgaagaagaccacta 227  
Db 2331 AATTAGAAATGCAAGCTGATGAATCCAAAAGTGAATGATGCTTGGAATGCTTCAA 2390  
Qy 228 ctccctcgcactcgtgtagagcatcagtcagactcatttacaagctgtaga----- 283  
Db 2391 ATGATTCAGATCCAGCTCATATAGAAAAGATCAACTCGTAGAGTGCCTGGATGAC 2450  
Qy 284 -----atccaccatctaagaagatcctcctgaagtc 321  
Db 2451 AAGCCACACAGAAAGCTTAGTACCAAAAGAGCTCGATGAAGTATCTCCAGCTT 2510  
Qy 322 cttctataaatttaagtttaacaagctgaatgagcttctggtctggtgagat 381  
Db 2511 CATTGAGAGATTATGAACTTAATTTAACTAAAGCCCTTATTTGTTGTTGTTGAT 2570  
Qy 382 tggctctgcttctaaatgagacgttcacagatatttccalcalttcgaaaaa 441  
Db 2571 TTTGTCCATTATTAATGAGGCGCTGCAACCATTTTGCATTAATTTTCAAAAGTTA 2630  
Qy 442 taacacatgttggaaataatgataaaacacatlaag---calgatgcagaattat 498  
Db 2631 TAGGGCTTTTACAGAAATGATGATCTGAAACAAACAGCAATATGATTAATCTGTTT 2690  
Qy 499 ccaltgatcgtcatttgggtgtatttgccttgcagatttcaatgagagat 558  
Db 2691 CACTATGTTTTCAGCCCTTGGAATTTCTTTTAAATTTTCTTCAGGGTTTCA 2750  
Qy 559 ttacagcagagcagaggaatttlaagatgagatlaagacacttgcctcaagca 618  
Db 2751 CATTGGCAAGCTGGAGAGTCTCACACAGCGCTCCGATACATGTTTCCGATCCA 2810  
Qy 619 tttatcagagatatgacctggttltgataaagaagaacagcagagagcttgaaca 678  
Db 2811 TGCTCAGACAGAGATGATGTTGGTTTCATGACCTTAACACACAGCAGATGACTA 2870  
Qy 679 caatattagcatagatatagacaatcagaaggaacaggttccaagattgagct 738  
Db 2871 CCAGGCTGCGCAATGATGCTGCTCAAGTTAAAGGGCTATGTTCCAGGCTCTGTAA 2930  
Qy 739 taacacaaatgacactaacacagagacttcaatcattccttataatgagatg 798  
Db 2931 TTACCCAGAAATATGCAATCTTGAGACAGAAATATATCTCTCATCTATGTTGGC 2990  
Qy 799 agatgacattcctgattcgtatgattgctccagacttccgtagacaggaattgaa 858  
Db 2991 AACTAACCTTTACTTACTTACCAATTTGACCAATTCATGACAGAGTGTGTTGAA 3050  
Qy 859 ccgagcaatgactgatttgcacaaagaatgaagcaaaccttaagatcgtgaaaga 918  
Db 3051 TGAATAATGTTGCTGCAACACACTGAAAGATTAAGAAAGAACTAGAAAGTCTGGAGA 3110  
Qy 919 tagcactgaagctttagaagaatlagctactatagtgatcaaaagaagaagact 978  
Db 3111 TCGCTACTGAAAGCAATTAAGAACTTCCGACCGTTGTTTCTTACTAGACAGCAAGT 3170  
Qy 979 tcgagcaaatglatgaagatgcttcagacttcaacacagaataactcgaagaagcac 1038  
Db 3231 ACATCTTTGGAATTTACATTTTCCCTTACCCAGGCAATATGATTTTCTATGCTGAT 3290  
Qy 1039 agatattggaagctgtatgcatcagcagatccttatatttctcctagagcag 1098  
Db 3291 TTTGACATATGATGCTAGAGTTTGCAGGTAACATAGAAACCTTTGAGGAAAGCAC 3230  
Qy 1099 ggttcgattggaagcatttaattcaagctgtagcgaatgagcccgagagcagatca 1158  
Db 3291 GTTTCGGTTTGGAGCACTTGTGAGCAATTAATCTATGAGCTTTGAGAGTGTCTGT 3350  
Qy 1159 tagttttacagcaattcatalgtagatgcatcgaaaacagctcgttctgagctc 1218

```

Db 3351 TAGATTTTACGCTGTTGTTGGTCCATGGCCGTGGGGAAGTCACTTATTGCTC 3410
OY 1219 ctgaatcctccaagaacacgcgggctgcgcaltctgttgccttltgaaagaac 1278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3411 CTGACTATGCCAAGCCCAAAATATCAGACGCCACATCATCATATGTAAGAAAAAACC 3470
OY 1279 caaatatagacagcgcagtcacaaaggaaagacagacacatgtbaaggaaattag 1338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3471 CTTTGATTTGACAGCTACAGCAGGAGGCTTAATGCCAACAACATTTGAAAGAAATGTCA 3530
OY 1339 agtttcagaagctcttcttctatccatgtcgcacagatgtttccatccctgcgt 1398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3531 CATTTGGTGAAGTGTATTCATATCCACCCGACCGACATCCCACTGGTTCCAGGAC 3590
OY 1399 tatccctcaagtattgagcaggaagacagatgtgtgtggagcagcggtgtgga 1458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3591 TGACCTTGAGAGTGAAGAGGCGACAGCTGTGGTGGGCGAGTGGCTGTGGGA 3650
OY 1459 aagcactctgttcaacttctgagagacttttagaccggtgcaagagacatgtctgt 1518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3651 AGACACAGCTGTCCAGCTCTTGAGCGGTTCTACAGACCTTGACGAGAAAGTGTGC 3710
OY 1519 ttgattgtgtgagtcacaaagaaatgtacagtgctgcgtccgtcccaatagcaatcg 1578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3711 TTGATGGCAAGAAATAAAGCGAGCATGTTCAGTGGCTCCAGACACACCTGGGCATCG 3770
OY 1579 ttctcaagaagcctgtgcttctcaactgcagacatgtctgagaacatcgccatagtgaca 1638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3771 TGTCCACGAGACCCCATCTCTGTTTACCTGACACATTTGCTGAAACATTCCTATGAGACACA 3830
OY 1639 acaagcgtgtgtgcatatagatagatcaagaagccgcaaatgtgacaatatcatt 1698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3831 ACAGCCGGGTGTGTACAGGAGATCGTAGGCGAGCAAGAGGCGCAACATATCATG 3890
OY 1699 ctttatgtgaagctcctctgagaatacaacacacagtttgagctggaagagcacagc 1758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3891 CCTTATCGAGTCACTGCTTAATAATATACATAAAGTAGAGACAAAGAAATCACTACG 3950
OY 1759 ttctgcgcgcagaaacaaagactagctattgcgaaggctcttctccaaaacccaaa 1818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3951 TCTCTGGGGCCAAACAAACGACATGGCATAGCTGCGCTGTTGTAGACAGCTCATTA 4010
OY 1819 ttttatgtgtgaggaagccctcaagcctcgataatgacatgagatgaagtggtgtgacg 1878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4011 TTTTGCTTTTGGATGAAACCCACGCTCAGCTGTGATACAGAAAGTGAAGGTGTCCAG 4070
OY 1879 atgccttgaataagcagcagcaggaagacatgacctagtgctcactcaagcgtctctg 1938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4071 AAGCCTGGACAAAGCCAGAGAGGCGCACCTGCAATTGTGCTCACCGGCTGTCCA 4130
OY 1939 caattcagaacgcagatttgaatgtgtctgcacaaatgtgaagaatgaagaagaa 1998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4131 CCATCCAGAAAGACAGACTTAATAGTGTGTTCAAGAAATGGCAGACTCAAGGAGCTGACA 4190
OY 1999 ctcatcaagagctctcgtgaaatcgagacataatatttaagtaagtaagtaagcagtcag 2058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4191 CGCATACGACACTGCTGGCAAGAAAGGCAATCTATTTTCAATGGTCAAGTGTCCAGGCTG 4250

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: SmithKline Beecham Corporation
Street: 709 Swedeland Road
City: King of Prussia
State: PA
Country:
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09-120-513
FILING DATE: 22-JUL-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4233 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Query Match      31.2%; Score 645; DB 3; Length 4233;
Best Local Similarity 58.0%; Pred. No. 1.3e-184;
Matches 1221; Conservative 0; Mismatches 845; Indels 39; Gaps 3;

OY 1 cgagcaaggtcgcagctcaatcgtgtgtagcacaccgacttctactatcgaagtgcag 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1746 CTAGAGGAAGGCGCGGACCCACCATGTGATAGCTACCGGCTTGTACAGTCCCAATGCTG 1805
OY 61 attgatgtgagccctaaagatgtagtgctgacgagaagaagacacatgtaactaa 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1806 ACGTATTTGCTGCTGTTTGTGTGGTGTGTCATGTGTGAGCAAGAAATCTATGAAGACTCA 1865
OY 121 tggcaaaacgaagctatatacttcaactgtgtgtcagcagagatactaaagaactgtag 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1866 TGAAGAGAGAGGCGATTTTACTTCAAACTGTGCATGACACAGACTAGAGAAATGAATTG 1925
OY 181 aa-cagatgagatcaatgacatattctactgaagaagaacacaaactactctctgcac 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1926 AACCGAAGAAATATCTTATGATCCCAAAAGTGACACTGCTCTGTGAGTGTGACTTCAG 1985
OY 240 tctgtgaagagatcaatgaagcagactcaatgtgacaagcgtgagaagatccccaactaa 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1986 AAAAATCAAAATCTCTTAAATGAAGATCAATTCGCAAGAAATGCCACAGAGCAAG 2045
OY 299 -----agagataagcttctcgaagctctc 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2046 ACCAGAGAGAGACTTAGTTGGAAGAGATGTGATGATGAATGCTTATGGTTTCT 2105
OY 325 tatataaaatttaagtaaaacaaagcctgaatgagccttltgtgtctgaggacatgag 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2106 TTTGGCAGATCTTAAGCTAAATATATTAGATGCGCCCTATTAGTGTGGCTGTACTTT 2165
OY 385 ctctgtcttaagtgaactgttcatcagtaattttcacaatctcttgaataaataa 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2166 GTGCTGTATTAATGAGGTGACATACACAGTGTGGCCATAGTGTTCATAAGATTGTAG 2225
OY 445 ccatgtt---ggaataatgataaaacacataaagaatgatgcagaatlttcca 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2226 GGGTTTTCATGAGAGCGACGACCATGAACCAACCAAGGAATTGTACTGTTTCC 2285

```

Oy 502 tgaatcgcacatttggtggtatgttgcttgacgttaatttcacgtcagggatattt 561  
 Db 2286 TTTCTTTCTGTCATGGGAATGATTTCTTTGTTACGTACTCTTTCAAGGCTTCACAT 2245  
 Oy 562 aagcgagagcaggggaaattttaacgaatgaatgaacattcgcttcacagcatgt 621  
 Db 2346 TTGGGAAACCTGGAGAGATCTCCACAGGAGCTCCGATACATGATGCTTCAATCCATGTC 2405  
 Oy 622 tatataagatatgtctgtgttgatgaataaagcaacagcagaggtgttacacaa 681  
 Db 2406 TCGGACAGGATTAATAGCTGGTTTGATGACCATAAAAACCACTGGCTCGCTACATACCA 2465  
 Oy 682 tatagccatagatatagcaacaattcaaggaacaaaggttccagagatggtgtcttaa 741  
 Db 2466 GGGTCCTGATGACGGCTTCTAATGTTAAAGGGCTATGGGCTCCAGGCTGTGTAAGTTA 2525  
 Oy 742 caacaaatgcacataacatgagacttcaatcatcattcccttatatatagatggaga 801  
 Db 2526 CCCAGATATGACAAACCTTGCGACAGGAATATCTATCTTATGTTATGCTGAGGAGC 2585  
 Oy 802 tgcattcccgatctcgagatgagatgctccagttacttgcctgacaggaatgtgaaccg 861  
 Db 2586 TTACACTTTACTTGTAGTAAATTAATACACTCTATTGCTTGGGTGGAATTAATGAATGA 2645  
 Oy 862 cagcaatgactggtatgtgcaacaagaataagcaaacacttaagcatgctggaagaatag 921  
 Db 2646 AACTGTTGCTGGTCAGACCTTGAGAGCAAGAAGAGCTAGAGATCTGAGGAAGATCG 2705  
 Oy 922 caactgaagcttggagataatacgtactatgttcatatacaagaaggaagacgtctcg 981  
 Db 2706 CTACAGAGCAATGTGAATAACCTTCGCACTGTGTCTTTGACTAGGAGGACAGAAATTTG 2765  
 Oy 982 agcaaatgtatgaagaatgcttcaagactcaacacagaataactcctgaaagaacacaga 1041  
 Db 2766 AAACATATGTATGCCACAGACTTGACAGTACCATACAGAAATGCTTTAAGAAAGCACACG 2825  
 Oy 1042 ttattgaaagctgtatgacattcagccatgcttataatatattgcctatgacagaggt 1101  
 Db 2826 TCTTTGGGATCACCTTTCGCTTCACCCAGGCGATGATTTATTTTCTATGCTGCTGTT 2885  
 Oy 1102 ttcgatttgaagcctatttaattcaagctggaagatgacccagagggcatgttcatag 1161  
 Db 2886 TCCGGTTGCGGTGCTACTTGTGTGGCAGAGAACATCATGACTTTGAATAATGTTATGTTG 2945  
 Oy 1162 ttcttactgaattgcataatgagactatgagccatcgcaaaagcgtcgtttgtcctctg 1221  
 Db 2946 TATTTTCTGCTGTGCTTGTGTCATGCGCATGGCAGAGGAATACCAAGTTTCATTCGCTCTG 3005  
 Oy 1222 aatatccaaagccaaatcggggctgcatctgttgccttltgtgaaagaaccaa 1281  
 Db 3006 ACTGAGGGAAGGCCAAAGTCCAGCATGCCATCATCAGATCATTTGAGAAATCCCGC 3065  
 Oy 1282 atatacagcagcgagtcagaagaaggaagaaagccagacacatgtgaagggaattagagt 1341  
 Db 3066 AGATTGACAGCTTACAGCAGGAGGCTTGAAGCCTAATTTGTTAAGAAATGTTGAAT 3125  
 Oy 1342 ttcgagaagctcttcttctatccatgctgcacagatgtttcatcctcctgtgcttat 1401  
 Db 3126 TTAATGAGATATTTTCACTATTCACCCGACCCGACCAATCCAGTGTTCAGGAGACTGA 3185  
 Oy 1402 ccctcagatgacgaggaagaacagtagcatattgttgggagcagcggtctgtgaaaa 1461  
 Db 3186 GCTTCGAGCTGAAGAGGGGCAACGCTTCGCTGTGGGAGAGAGTGTGGGAGAGA 3245  
 Oy 1462 gcaacttctgtcaacttctgacagacttatagacccgctgcaaggaaagtgtcgtttg 1521  
 Db 3246 GTACAGTGTGTCCAGTGTCCAGGCTTCTACAAACCCATGGCTGGAACATGTTTCTAG 3305  
 Oy 1522 atgtgtgtgagcaaaagaattgaatgacagtggtcgttcccaaatagcaatcgttc 1581  
 Db 3306 ATGGCAAGAAATAAACACTCAACGTCAGTGTGCTCCGGCCCACTGGGCATTTGTT 3365

Oy 1582 ctcaagagcctgtgctcttcaacatgacatgtctgagaacatcgcctatgtgtacaaca 1641  
 Db 3366 CCAGAGAGCCCATCTCTGTTTGACTGCACGATCACCCAGAAACATCGCTACGGAGACACA 3425  
 Oy 1642 gccgtgtgtgcatatagatgaataaagacgcaaatgagcaaaatcatcattt 1701  
 Db 3426 GCGGTCTGTTGTCTCATGAGAGATGTTGAGGGCCGCCAGGGAGGCCAATCACCAGT 3485  
 Oy 1702 ttatgaagctctccctgagaataacacacaaagtltgagctgaagaagacacagctt 1761  
 Db 3486 TCATGCACCTACCTGCTGAGAAATACACACCAAGTGTGGAGCAAGAGGACTCAGCTGT 3545  
 Oy 1762 ctgagcgccaagaacaaagactagctatgcaaggtctcttccaaaaccacaaatt 1821  
 Db 3546 CGGCGGGCAGAGCAGCGCATCGCATCGCGCGCTCTGTCAACACCTCACAATCT 3605  
 Oy 1822 tatgttgaatgagcacttcagccctcagataatgacagtgaagaagtgttcaagatg 1881  
 Db 3606 TACTTCTGATGAAGCAGCATCATGCTGTGATACGGAGATGAAGATGATGCTCAGGAAG 3665  
 Oy 1882 cccttgaataagccagagacggaagagacatgctcagtagtgcactcaaggtctcgtcaa 1941  
 Db 3666 CGCTGACAAAGCCAGGAGGAGCGCCGACCTGATTTGATCGGCAACCGCTGTCCACA 3725  
 Oy 1942 ttcagaacgagatltgatatgtgttctgcacaaatggaagaataaaggaagaagactc 2001  
 Db 3726 TCCAGACGCAAGCTGATGTATGTGATTCAGAACGCGCGATCAGAGACAGGACACC 3785  
 Oy 2002 atcaagagcctcctgaagaatcagacatatatttaagttagtgaatgacagtcagtc 2061  
 Db 3786 ACCAGCAGCTGTGTGCGCCAGAAAGGCAATATTTCTGATGTGTTCAAGCTGAGCAAGC 3845  
 Oy 2062 agtga 2066  
 Db 3846 GCTCA 3850

RESULT 11  
 US-09-450-105-1  
 ; Sequence 1, Application US/09450105  
 ; Patent No. 6169166  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kimberly Anne Brun  
 ; APPLICANT: Richard James Chenery  
 ; APPLICANT: Harma Eileus  
 ; APPLICANT: John Anthony Feild  
 ; APPLICANT: Lin yue  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
 ; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF  
 ; FILE REFERENCE: GP-50008-D1  
 ; CURRENT APPLICATION NUMBER: US/09/450,105  
 ; CURRENT FILING DATE: 1999-11-29  
 ; EARLIER APPLICATION NUMBER: 09/120,513  
 ; EARLIER FILING DATE: 1998-07-22  
 ; NUMBER OF SEQ. ID NOS: 2  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ. ID NO 1  
 ; LENGTH: 4233  
 ; TYPE: DNA  
 ; ORGANISM: HOMO SAPIENS  
 US-09-450-105-1

Query Match 31.2%; Score 645; DB 4; Length 4233;  
 Best Local Similarity 58.0%; Pred. No. 1,3e-184;  
 Matches 1221; Conservative 0; Mismatches 845; Indels 39; Gaps 3;

Oy 1 cgagcaaggtcgcagactcaatcgtgtgtagcacaccgacttctactatttgaaagtgcag 60  
 Db 1746 ctgaggaagcgcggaaccacatltgtatagctcacccgttctacagltgcacatgtcgt 1805  
 Oy 61 attgatgtgacacctaaagatgaaatgctggtgcggaagaagagacacatgctgaactaa 120

Db 1806 acgtcatctgctgtttcttgatggttggtgcatctgtgtgagcaagaatacatgaagatctca 1865  
Oy 121 tggcaaacgaggtctatatatttcaacttgtgatgtccaagatattataaaagctgag 180  
Db 1866 tgaagagaagggcalttacttcaaaactgtgcatgacacagactagaggaaatgaatgtg 1925  
Oy 181 aa-cagatgagatgcatacttcaatcttgaagaagaacccaactcaatctctctgcac 239  
Db 1926 aaccgggaatataatgtctcatgaatcccaagtgacactgtgtccctcgtgattctcg 1985  
Oy 240 tctgtgaagagcatcaagtcagacttcatgtgcaagctgtgaggaatccccaactca- 298  
Db 1986 aaaaatcaaatctccttcaatgaagagatcaattcgcagaagatcatccagacaagaac 2045  
Oy 299 -----agagataagttctcctgaagttctc 324  
Db 2046 accaagagaagaagacttactgtcgaagaagatgtgattgaagatgctcctaagttctc 2105  
Oy 325 tatataaattttaagtttaaaccaagcctgaaatgacctgtgtgtctgtgggacaattg 384  
Db 2106 ttgcaagatccctaagctaaatataatagtgaatggcccaattagttgtggtgtgatact 2165  
Oy 385 ctctgtctcaaatggaactgtgtccacglatcttccatcatcttgcgaataatataa 444  
Db 2166 gtgctgtataaattggtgtgcatacaaccagttgttgcatagtgttttcaaatgttag 2225  
Oy 445 ccaattt---ggaataatgaataaacacatlaaagcatgacagagaattatcca 501  
Db 2226 gggatttctcaagagacagacgacccaatgaacccaacgaatgtlaactgtttccc 2285  
Oy 502 tgaattctgcatcttgggtgttattgttctgtcagttatcttcaagggatatttt 561  
Db 2286 tctcttctcgtcatagtgaaatgatttcttctgttaagtaacttcttcaagccttaac 2345  
Oy 562 acggcagagcaggggaaattttaacgagtgaataagacactgtgcctcaagccatgt 621  
Db 2346 ttgcaagagctgagagatccctcaccaagcgactccgatacatgtcttcaatccatgc 2405  
Oy 622 tatatcagatatttgcctgtgttgatgaagaagaaacgcacagagcttgcacaa 681  
Db 2406 tgcagacagatataagctgtgttgaatgacataaaacacacactgtgcctgactacca 2465  
Oy 682 tatctgcatagatatagacaacatccaaggaacaggttccaagatgcgtcttca 741  
Db 2466 ggtctgcatagtgacaccttcaatgttaagggtcctatggtcccaagctgtcgtcatga 2525  
Oy 742 cacaataatgcaactaacatgagacttcaatcatcttcaatataatgagtgagaga 801  
Db 2526 ccaggaatgtgacaacctgtgcacaggaattatctatccttagtctatgtcgtgcagc 2585  
Oy 802 tgaacttctgattctgagttatgtctcagttatctgtcgtgacaggaatgtgaaccg 861  
Db 2586 ttaacctttaaacttgaatataccactcatctgttctgtggtgaattatgaatga 2645  
Oy 862 cagcaatgagctgattgtgcacaagaagataagcaagaacttgaagctgtgtgaaatag 921  
Db 2646 aactgtgtcgtgccaagccttgaaggaacaaagagcttagagatcctcgggaatag 2705  
Oy 922 caactgaagcttggagaatatacgttaccatagtgtcatlaacaagggaaaaagccttg 981  
Db 2706 ctacagaagcaattgaanaactccgcactgtgtccttctgactagaggcagaatttg 2765  
Oy 982 agcaaatgtatagaagatgtctcagactcaacacagaataatcctcgaagaagaacaga 1041  
Db 2766 aaactatgtatgccagagctgtgcagatcacatacagaatgtcttgaagaagcacag 2825  
Oy 1042 ttattggaagcgtgtatagctcagcagccttataattttgcttcagcagaggt 1101  
Db 2826 tcttgggatacactcgtcctccacacaggaagcatgattatcttccatgtcgtctgt 2885  
Oy 1102 ttgcatttggagccttatttaattcaagctgtgacgaatgacccacagaggcatgtcatag 1161  
Db 2886 tccggtcgtgctactactgtgtgcacaggaactcatgacgttctgaataatgtatgttg 2945

Oy 1162 ttttacttcaattgcatatgagctatgtgccactcgaaaaacgctcgttttgcctctg 1221  
Db 2946 tatcttgcgtgtgttcttctgtgtgcatagtgcagcaaggatatacagttatctgcctcg 3005  
Oy 1222 aatattccaagaagccaatcggggctgcgcactctgttttgcctgtgtgaaagaaccaa 1281  
Db 3006 actacgcaaggccaagatcagcatcccaactcaatcagatcatcttggaataatcccg 3065  
Oy 1282 atatacagcgcagtcagtcagaagaagaaaaagccagacacatgttgaaggaaattgagt 1341  
Db 3066 agattacagctacagcagcagggcgttgaagccttaattgttgaaggaaatgtgaaat 3125  
Oy 1342 ttcggaagctccttcttcttctatccatgctgcgccagatgtttcaaccgcggtttat 1401  
Db 3126 ttaattgagttcatgttcaacttccaccgcgaacacatccacagtgtcttcaaggactga 3185  
Oy 1402 ccttcagtatgagcgaagaagacagtagcatattgtgggagcagcgtgtggaanaa 1461  
Db 3186 gcttcgagtgagaagaagggccaacgcttcgcctgtgtggcagcagtggtcgtgggaaga 3245  
Oy 1462 gaacttctgtcaactctgcgagagacttatagacccgtgcaaggaacagtgtcttg 1521  
Db 3246 gtaacagtgtccagctgtcgcagcgtcttatacaaccccatgtgtgaaacagtgttctag 3305  
Oy 1522 atgtgtgatatcaanaagatgaatgatacagtggtgtccgttuccaataagcaatcgttc 1581  
Db 3306 atggtcaaaanaataaaacacatcaacgctcagtggtccgcgcacccgtggatctgt 3365  
Oy 1582 ctcaagagcctgtgtccttcaactcagcagcatgtctgagacaatcgcatagtgtgaca 1641  
Db 3366 ccagagagccaatccctgttgaactgtgacatcacccagaacatcgcctacggaacaa 3425  
Oy 1642 gccgtgtgtgccatagatgagatacaagaagcgcgaatgacgacaataatcattct 1701  
Db 3426 gccgtgtgtgttcatagagagatcgttgaaggccgcagggagagcacaatccacagt 3485  
Oy 1702 ttattgaagttcccttcggaataatacaacacaagaattgagctgaagaagcagctt 1761  
Db 3486 tcatcgaactcaactgtcctggaataatacaacacacagatgtggagacaaaggaactca 3545  
Oy 1762 ctggtgcgcaagaacaaactagactatgtcaaggctccttctccaanaacccaatatt 1821  
Db 3546 cgggtgcggaagaagcagcgcactgcacatgcgcgtcgcctcgtcgaacagcctcacatct 3605  
Oy 1822 tattgtgataagagcacttcaagcctcogataatgaacagtgaaggtgtgttcacatg 1881  
Db 3606 tactctgatatgaagcagacatcagctctgatacgagagatgaaaggtcgtccaggaag 3665  
Oy 1882 cccctgataaagcagaagcgggaagacatgcttaagtgttcaactcaagcttctgcga 1941  
Db 3666 cgtctgacaagaagcagaagagcgcgacactgtcatctgtgacgcgcacgcctgtccaca 3725  
Oy 1942 ttcaagacagagattgatagtgttctgcacatgtgaaagatacaaggaacagaagatc 2001  
Db 3726 tccagaacagagacttgcattcgtgtgtatccaaagccgaaggaaggaaggaaccc 3785  
Oy 2002 atcaagagctcctgagaanaatcagacatatatttaagttagtgaatgcaagtcagtc 2061  
Db 3786 accagcagctgtgcgccagaagaagcatctatttctcagatgtgtcagctgtgagcaagc 3845  
Oy 2062 agtga 2066  
Db 3846 gttca 3850

RESULT 12  
US-08-612-734B-1  
; Sequence 1, Application US/08612734B  
; Patent NO. 5914246  
; GENERAL INFORMATION:  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Skatrud, Paul L.





Db 3670 GGCAGCAGGAGGCGATGTTCTGTGGCGCCAAAGCAAGCTGTGGCATTTGCCGAGCC 3729  
QY 1800 cttcccaaaaaccccaatttatttggatgagccacttaagccctcgtataatgac 1859  
Db 3730 CTCTTGGGGATTCCTCAAAATCCTTCTCTCGATGAAGGAGCGTCAAGCCCTGACTCCGAG 3789  
QY 1860 agtgaagaagtgtgttcagcatcgtccttgataaagccaggaagggaagacatgctagt 1919  
Db 3790 TCAGAAAAGGTCTCCAGGCGGCTTTGGATGCCCTGCCAGGCGCAGACCAATGCC 3849  
QY 1920 gtacactacaggtcctctcgtcaatcagaaacagagattgtagtctcgtcacaatgga 1979  
Db 3850 GTTGACACACCGACTCAGCAGCATTCAAAAGCGGACGCTTATCTATGTTTCAGCAAGGC 3909  
QY 1980 aagtaagaagacaagaacatcacaagagctccttgagaatcagacatatatttaag 2039  
Db 3910 AAGATCGTCCGAAGCGACGACGACGACTGTCCAGAAAAGGCGGCTACTACAG 3969  
QY 2040 ttaagtaatgcacag 2054  
Db 3970 CTGCTCAACTTCAG 3984

RESULT 14  
US-08-996-545-3  
Sequence 3, Application US/08996545  
Patent No. 5928898  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,545  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-996-545-3  
Query Match 13.7%; Score 283.8; DB 2; Length 4002;  
Best local Similarity 35.2%; Pred. No. 1.7e-75;

Matches 555; Conservative 213; Mismatches 807; Indels 0; Gaps 0;  
QY 480 catgatcagaattattcatgatattgctcatttggtgtgtatttcttgatg 539  
Db 2410 CAUGAUGCGGAUUCUGGUCUAGUGGUCUUCUGGUGUGUACAUUCAGUUAUCAG 2469  
QY 540 tattctgcagggattattttaaagcagagcaaggggaatttaaagatgagataa 599  
Db 2470 CAGUACACAAUUGUGUGUGAUAUUGCCGUAUGUCUGGAGAGACUUAUUGCGCGGAGA 2529  
QY 600 caattgaccttaaaagccatgtatcagagatttccctggttgatgataaagaaac 659  
Db 2530 AGCACUCCUUCUGAGAGUADUCCGUCACAAACAUUCUUCUUGACAGGAAGAAAU 2569  
QY 660 agcacagaggtcgtgacaacatataggcctagatataagacaattcaaggagacaa 719  
Db 2590 AGCACCGGCGCUCGACUUCUUCUUGCCACCGAGACAGAAAGCAUUCUCCGUGUAGC 2649  
QY 720 ggtccagagattgctgtcttaacacaaatgcaactacatagggacttcagtatacat 779  
Db 2650 GGUUGAUCUUGAGCAGAUUCUGAUGACUCCACAGACCUAGAGGCGCUAUAUAU 2709  
QY 780 tccttataatgagatgagagatgacattcctgattctgattgtcagatcagtcgcc 839  
Db 2710 GCCUGGCGAUGGUGGAAUUGGCGUUAUUGUUGUUCUGGUGUGCGGUCUCCUG 2769  
QY 840 gtgacaggaatgattgaaccgcagcaatgactgatttgccacaagaatgaagcaaga 899  
Db 2770 GCAUGCGGUUCUACCGAUUUAUUGUACCCAGUUAUUCAGUCCUCCAGACUUCGU 2829  
QY 900 ctaagcatcgtggaagatagcacatgaaagctttgagaatactactatagtgtca 959  
Db 2830 UAUGAGGUAUUGCAACUUAUUGUGGAGGCUAUCUUAUCCGACAGUUGUGUCA 2889  
QY 960 ttaacaagggaagaaagccttcagagcaatgtatgaagagatgctcagactcaacaga 1019  
Db 2890 UUAACCCGGAAGAGAUUGUGGAGAUUACCAUUGCCAGCUUAGCGCAAGAGAGG 2949  
QY 1020 aataccctgaagaagacagattattggaagcgtgatgatcagcagtccttata 1079  
Db 2950 ACCAGUUAUUCUGUUGUAGGUAUUGUUAUUGGUGUGGUGCGACGACUUGU 3009  
QY 1080 tatttgcctatgcagaggttcgaatttgagacttaattcaagtgcagatg 1139  
Db 3010 UUCUUCUGGUGGUGGUGUUGUUGGAGGAGGACACUUCUUGGACACAGAGUAU 3069  
QY 1140 accccaagagatgattcatttctacgtcaattgcatactgagctatgacatgga 1199  
Db 3070 GACAUAUUCGCCUUCUUGUUGUUGUUCGAGAUUCUUCUUGGUCUCAAUCCGCGGC 3129  
QY 1200 aaagcgtcgttttggtcctgaatttcacaaagcaaatcgggggtgcacatgttt 1259  
Db 3130 ACCGUUCUUCUUGCACCAGACUUGGCGAAGGCGAAGAUUGGCGCGCGGAUUCGA 3189  
QY 1260 gcttgttgtaaaagaaacaaatataagacgcgcagtcagaagaaggaaagccagac 1319  
Db 3190 CGACUUGUUGACCGAAAGCCAAAUUGAUAUUGUUGAAGAGGCGAGACUUGGAA 3249  
QY 1320 acatgtgaaggaaattagatgattcagagaagtccttctctcatcatgctgcagat 1379  
Db 3250 ACGGUGAAGGUAUUGAUAUUGAUAUUGAUAUUGAUAUUGAUAUUGAUAUUGAUAU 3309  
QY 1380 gtttaaccctcgtgtgttccctcagatgtatgagcagaaagacagtagcatgttg 1439  
Db 3310 CAGCUGGUGCGGCGGUGUGACCUAGCGUAGGACCUAGCAUAUUGGUGGUGUUC 3369  
QY 1440 gggagcagcgctgtggaagaagcaatctgttaactctgcagagacttatgacccc 1499  
Db 3370 GGACCCAGCGGUGUGGAGAAAGUACCAUUGCAUUGCUUGAAGCGUUAUUGAGUGCG 3429  
QY 1500 gtgcaagacaagtgtgttgatgtgtgtgtaagaaagaaattgaatgtacagtgctc 1559  
Db 3430 AUUGCGGCGUCCAUUCUUGUUGAUGGAAAGCAUAUAUAUAUAUAUAUAUAUAUAUA 3489



```

QY 1560 cgtcccaatagcaatcgtctcctaagcagcgtgtccttcaactgcagatgctgag 1619
    ||| : ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 3490 CGCAGCUUUCUGUCACUGUACGCCAGCAGCCGACACUGUACAGGACCAUACAGGAA 3549
QY 1620 aacatgcctatgtagcaacagcgcgtgtgtgcattagatgagatcaagaagcgca 1679
    ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 3550 AACAVUCUACUGUACUGUACGAGAGUACGACCGAAGAAUUCUUGAUUUAAGCUUGC 3609
QY 1680 aatgcagaatataatcattcttatttgaagtctccttggaataatacaacaagt 1739
    ||| ||| ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 3610 AAGAGCGUAUUCUACUACGACUACUACUGUCGCCGAGGCGCUUUAUACACUUDU 3669
QY 1740 ggaactgaagagcagcgtcttcctggccgaacaaagactagctattgcaaggct 1799
    ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3670 GGCAGCAAGGAGGCAUGUUCUGCGCGCCCAAAAGCAACUGUGCCCAUUGCCGAGCC 3729
QY 1800 ctctcccaaaacccaaatttattgtgtgtagtgagcacttcaagccctcgataatgac 1859
    ||| : ||| ||| ||| : ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 3730 CUUCUUCGCGGAAUCCCAAAUUCUUCUUCUGAUAGCAGGUCGACCCUGACUCGAG 3789
QY 1860 agtgaagatgtgtcagcatgacctgtataaagcgaagcggaagacatgctagt 1919
    ||| ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 3790 UCAGAAAAAGGUGUCAGGCGCUUCGAGUCCGUGCCGAGCGCAACAUAUCCGCC 3849
QY 1920 gtcaactcaagctctctgcaatcagaagcagatgtgtagtgtgtctgcacatgga 1979
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3850 GUUCGACACCGACACGACGACUACAAAGCGGAGUACUACUACUUCUUCGACCAAGCC 3909
QY 1980 aagaataagagacaagacatcacaagaagctctgagaataatcgacataattttag 2039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3910 AAGUUCGUGCAAGCGGAGCGACGCAACUUGUCCAGAAAGGCGGUGACUACGAG 3969
QY 2040 ttagtgaatgcag 2054
    ||| ||| |||
Db 3970 CUGGUCACUUGCAG 3984
    ||| ||| |||

```

# RESULT 15

```

US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
;
US-09-328-320-1

Query Match 13.7%; Score 283.8; DB 4; Length 4002;
Best Local Similarity 48.8%; Pred. No. 1.7e-75;
Matches 768; Conservative 0; Mismatches 807; Indels 0; Gaps 0;

QY 480 catgatgcagaattatcctatcagatgattgcatttgggtgtatttgccttgcagt 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2410 CATGATGCGGATTTCTGGTCATTGATGTCTTGCTTGTTGATCATTCATTATTCACG 2469
QY 540 tattcatgcaggaattattttagcggcagagcagggaattttagatgattaga 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2470 CAGTCAACCAATGATGCTGATTCCTCCGATCTCCGAGACATTAATGTCGCCGAGA 2529
QY 600 caacttgccttaaaagccatgtctatatacagatattgccttggttgaagaagaaac 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2530 AGCACTGCTTTTCGACGACATCTCCGTCGAAGCATTTCTTTTTCACAGAAAGAAAT 2589
QY 660 agcagaagagcttgacacaatatagccatagatatagacaataatcaaggacaaca 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2590 AGCACCAGGCGCTGACACTCTTCTGTCACCGAGACGACGACATCTCCGCTGTACG 2649
QY 720 ggtccagagatgtgcgtcttaacacaaaatgcacataatgagcttcaattacat 779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2650 GGTGTGACTCTAGGACGACGATTTGATGACCTCCAGACCTAGAGAGCGCTATCATTAAT 2709
QY 780 tccctataatagatggagatgagatcctcgtatctgagatagctcagctactgac 839
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2710 GCCCTGGCGATTGGGTGAAATTTGGCTTAGTTGATTCGCTGTGCGGTTCTCTCG 2769
QY 840 gtacagaagaatgatgaaacgcgaagaatgagatgttgcacaacaagaataagaaga 899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2770 GCATGCGGTTCTACCGATTTATATGCTAGCCAGTTTCATWCAAGCTCCAGCTTGTCT 2829
QY 900 cttaagcatgtctgaaagaatagcaacttgaagcttggagaataatacgtactagtgtca 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2830 TATGAGGATCTGCAAACTTCTGCGAGGCTACATCTCTATCCGACAGATTGGGTCA 2889
QY 960 ttaacaaggaaaaagcccttgagaacaatgltgagaagatcttcaagctcaacaga 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2890 TTAACCCGGGAAGGATGTCTGGAGATTTACATGCCACCTTGACCCACAAAGGACAG 2949
QY 1020 aataccctgaagaagacacagatattggaagctgttatgcatcagcagatgcttata 1079
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2950 ACCAGTCTAATCTGTCTTGAGGTATCCCTGTATATAGCCGTGCGGACGACATCTGTT 3009
QY 1080 tatttgcctatgcagcagggttctgatttgaagcctatttcaatcaagctgagaagat 1139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3010 TTCTTCTGCGTTGCGCTCGGGTTTGTGTAAGGAGGACACTTCTGTGTACACAGATAT 3069
QY 1140 acccagagaggagatgtcaatgatttgaatgcaatgcaatgcaatgcaatgcaatgca 1199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3070 GACATTTTCCGCTTCTTTGTTGTTCTTCGAGATTTCTTTGGTGCTCAATCCGCGGCC 3129
QY 1200 aaaacgctgttlttgctctcgtatataatccaaagcacaatcgaggcgctgcatctgtt 1259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3130 ACCGCTTTTCTTCTTTCACACACATGAGGCAAGGGAAGAAATGCGGCCGCAATTTCCGA 3189

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 15:53:32 ; Search time 866.72 Seconds  
(without alignments)  
1997.512 Million cell updates/sec

Title: US-09-873-409-9

Perfect score: 2066

Sequence: 1 cgagcaaaagtcgactacta.....atgcacagtcacgtgcagtga 2066

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	739.8	35.8	3924	21	AAZ94742
2	739.8	35.8	3924	21	AAZ88974
3	712.6	34.5	4669	14	AAO52726
4	712.4	34.5	2726	15	AAO70907
5	712.4	34.5	2726	15	AAO70916
6	712.4	34.5	2726	18	AAAT3322
7	712.4	34.5	4349	22	AAH57442
8	712.4	34.5	4646	15	AAO72872
9	712.4	34.5	4646	21	AAZ94738
10	712.4	34.5	4669	8	AAAT70752
11	712.4	34.5	6505	17	AAAT13394

12	710.8	34.4	3860	21	AAZ49332
13	710.8	34.4	3860	21	AAZ49333
14	710.8	34.4	3988	21	AAZ88973
15	710.8	34.4	4186	22	AAAF86127
16	710.8	34.4	4195	22	AAAF86128
17	710.8	34.4	4264	19	AAV66533
18	710.8	34.4	4264	19	AAV66534
19	710.8	34.4	8630	21	AAZ24042
20	709.2	34.3	4378	11	AAQ04522
21	709.2	34.3	8630	21	AAZ24041
22	704.4	34.1	4669	19	AAV32645
23	702.4	34.0	4279	22	AAO34488
24	702.4	34.0	4279	22	AAO34504
25	702.4	34.0	4279	22	AAO3505
26	700.8	33.9	4279	22	AAO3506
27	700.8	33.9	4317	22	AAO34489
28	658	31.8	4369	21	AAZ52047
29	658	31.8	4425	21	AAZ52048
30	646.2	31.3	4189	21	AAZ49334
31	646.2	31.3	4313	14	AAO38950
32	645	31.2	4233	21	AAZ90198
33	645	31.2	4233	22	AAAF27498
34	637.2	30.8	4788	21	AAZ49335
35	490	23.7	4776	21	AAZ94744
36	309.4	15.0	386	22	AAH94091
37	303.4	14.7	3512	20	AAV69395
38	303.4	14.7	4175	20	AAV69392
39	303	14.7	2698	20	AAV69394
40	297	14.4	2681	20	AAV69396
41	285.8	13.8	4047	20	AAV60201
42	283.8	13.7	4002	20	AAV85823
43	271	13.1	4051	21	AAO9019
44	268.2	13.0	4102	21	AAO45942
45	260.8	12.6	1957	21	AAO35415

#### ALIGNMENTS

RESULT 1	
AAZ94742	
ID	AAZ94742 standard; cDNA; 3924 BP.
XX	
XX	AAZ94742;
AC	
XX	
XX	01-AUG-2000 (first entry)
XX	
XX	Human ATP binding cassette ABCB4 (MDR3) cDNA.
DE	
KW	ABCB4; ATP binding cassette; human; cholesterol; lipid disorder;
KW	atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW	lupus erythematosus; diagnosis; gene therapy; MDR3;
KW	multidrug resistance; chromosome 7q21; ss.
XX	
OS	Homo sapiens.
XX	
PN	W0200018912-A2.
XX	
PD	06-APR-2000.
XX	
XX	21-SEP-1999; 99MO-EP06991.
XX	
XX	25-SEP-1998; 98US-0101706.
PR	
XX	(FARB ) BAYER AG.
PA	
XX	
XX	Schmitz G, Klucken J;
PI	
XX	WPI; 2000-293151/25.
DR	
XX	
XX	Adenosine triphosphate binding proteins useful for identifying agents
PT	for treating atherosclerosis and other inflammatory disorders -
XX	

PS Claim 9; Page 117-119; 154bp; English.

XX The present sequence is that of human ATP binding cassette  
CC subfamily B protein ABCB4 cDNA. The cDNA was identified using a  
CC differential display method in which monocytes from peripheral  
CC blood were subjected to macrophage differentiation and cholesterol  
CC loading with acetylated low density lipoproteins and subsequent  
CC degrading with high density lipoprotein (HDL) to identify  
CC cholesterol sensitive genes. The gene maps to chromosome 7q21  
CC and is also termed MDR3 (multidrug resistance). The invention  
CC provides cholesterol sensitive ABC genes (see AA294734-63). These  
CC genes, and polypeptides encoded by them, can be used for diagnostic  
CC and therapeutic applications, and for biochemical or cell-based  
CC assays to screen for pharmacologically active modulator compounds  
CC useful for the treatment of lipid disorders, atherosclerosis or  
CC other inflammatory diseases such as psoriasis and lupus  
CC erythematosus.

XX Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Query Match 35 8%; Score 739.8; DB 21; Length 3924;

Best Local Similarity 60.6%; Pred. No. 7.9e-189; Indels 33; Gaps 2;

Matches 1268; Conservative 0; Mismatches 792; Indels 33; Gaps 2;

1 cgaagcaagctcgactacatcgtgtgaagacacacgactcttactatcgaagtgcag 60  
1765 ccagagaaagccggagccacatctgtgtatgacacacgactgtccagctcgaatgcag 1824  
61 attgattgtgacccctaaagatggaatgctgctgcgagaaagagacatgctgaactaa 120  
1825 atgtcatcgtcgtggttgcgtatgaaatgtgtgagcagaagacacagcgaactga 1884  
121 tggcaaaacagagctctattatctactgtgat----- 154  
1885 tgaagaaagaaaggggtgacttcaactgttcaacatgcagacatcagaagccagatcc 1944  
155 -gtcacagagatataaagctgtatgaacagatgagatgacatgacatctctaa 213  
1945 agtcaagaagatgtgaactaataatgtaaaagctgcacactagaatggtcccaaatgtgc 2004  
214 gaaagcaactcaactcctctgcaactcgtgtgaagagacatcaagcgaactcttaca 273  
2005 ggaatctcgcctatttgaagcatctcaatcagaanaaaccttaaaatctacaatgtgtc 2064  
274 aggc-----tgaagaaatccacccaatctaaagataagagctccgaagctcctat 327  
2065 agaagagccttgatgtggaacacgaatggaacttgaagcaaatgtgcacacgaatgtccttc 2124  
328 taaaattttaaagttaaacaagccttgaatgcctttgtgtctgtggaacatgtgctt 387  
2125 tgaagagctcgaactgaataaacaagaatgagccctactctgtcgtggaacagatgtg 2184  
388 ctgttctaataatggaactgttcatccagatattccatcatcttgcgaanaattatacca 447  
2185 ccattgcgaatgggggctctcagcgcatttccatcatatctcagagatctacagca 2244  
448 ttttggaaataatgataaaacacatgaagcagtatgcgaanaattattccatgatat 507  
2245 ttttggacagcgatgagtcagtgaaagcgagaaagtgcaacatctccttgcatt 2304  
508 tcgcatatttggatcttgcattgtctgcagtatttcatcagaggaattattttagcga 567  
2305 tctatttctgggaattattcttcttacttctcctcctcagagtttccagtttggga 2364  
568 gagcaggggaaattttaacgaatgaatgaacacatgtgccttaaacgaatgtatatac 627  
2365 aagctgtgagatcctccacagaagactgcgttcaatgtctttaaagcaatgtctaaagac 2424  
628 aggtatgtcctgtgtgtatgataaagaaacagcagcggtgttgaacaacatataag 687  
2425 aggcacatgactggttltgatgacataaaacagtaactgtgtgacatttctcaagaacttg 2484

688 ccatagatataagcacaaatcaagsgcaacaggtccaggaattgctgtcttaacacaa 747  
2485 ccacagatgtgcgccaaagctccacagagccacagaaacaggttgggttcaattatgcacaa 2544  
748 atgcaactaacatggaacttgcagttcatcttccattatataatgagtggagatgacat 807  
2545 atatacctaacttggaaactggaactgatatatacaattatctcagtttgggttaacc 2604  
808 tctgtattcgaatgtatgttccagacttgcgttgacaggaatgatgaacccagcaaa 867  
2605 tattgtcattagcagttgttcccaattatgtcgttgcgaagaattgttgaatgaattgt 2664  
868 tgcactgatttccacacaaagaataagcaagaacttaagcattgtctggaagaatgcgaacg 927  
2665 tggctggaatgtccaaaagataaaagaacttggaaagctgtctggaagaatgtgcacag 2724  
928 aagcttggagaataacgactagttgcatcttaacaaaggaanaagccttcagcaaa 987  
2725 agcgaatagaataatataagacagttgtgtccttgaacccaggaagaanaattgaaatcaa 2784  
988 tgtatgaagatgtctcagaactcaacacagaataatcctcgaagaagacacagatattg 1047  
2785 tgtatgttgaanaaatgttatgagaccttaacaggaattctgtgcagaagcacaatctatg 2844  
1048 gaagctgttatgcatcagccatgcttataatttgcctatgcagcaggttccgat 1107  
2845 gaattactttagatctccacaaagcatattgtatttctcctatgcggtgtgttccgat 2904  
1108 ttgagccttatttaattcaagctggagaaatgaacccagagggcagtgcttaagttttaa 1167  
2905 ttgtgtcatatcloatgtatgtgaatgacataatgcctcagaagatgtatctcgtgttct 2964  
1168 ctgcaattgcatatgagactatgccaatgcgaanaaacgctcgttcttgcctcctgaatat 1227  
2965 ctgcaattgcatatgagactatgccaatgcgaanaaacgctcgttcttgcctcctgaatat 3024  
1228 ccaagccaaatcgggggtcgtgcacatgttctcgttgcgttgcgaanaaaccaatctag 1287  
3025 cttaagatgaagctgtctgcgcgcacattatcctatgcgttcttgaanaaacacctctgatt 3084  
1288 acagcgcagatcgaagaaggaagaaagccagacacatgttgaaggaattatgagtttcag 1347  
3085 acagcgcagatcgaagaaggaagaaagccagacacatgttgaaggaattatgagtttcag 3144  
1348 aagctcttcttcttccatctgtcgcgaatgttctcctcctcgttgcgttctcctca 1407  
3145 aagctcgttcttccatctgtcgcgaatgttctcctcctcgttgcgttctcctca 3204  
1408 gtattgagcgaagaagaacagatgacatttgtggtggaacagcgtgtggtggaagaacacatc 1467  
3205 agtgaagaaggaagcgaacacatgacatttgtggtggaacagcgtgtggtggaagaacacatc 3264  
1468 ctgttcaactctcagagacttataagcccggaaggaaggaacagtgctgttgcattg 1527  
3265 tggccagctccctggagcgtgttctcagaccccttggcggggagacagtgcttctcgtatg 3324  
1528 tgaatgcanaaagaattgaattgacatgctcgttcccaaatagcaatcgttcccaag 1587  
3325 aagaagaacaaagaacatcaatgttccagtgctcagagctcaacatcgtatcgtcag 3384  
1588 agcctgtccttccaaactgcagacatgtctgaagaacatcgtcgtatgtaacacagcgtg 1647  
3385 agcctatcctatttgcagtcagcattgcgggaatattgtccttgaagaacacagcggg 3444  
1648 tgggtcattatgagatgagatcaagaagccggaanaatgcgaacaaatatacatcttattg 1707  
3445 ttgtatccagagatgaattgtgagtgcaagcaagctgcgaacaaatatacatcttattg 3504  
1708 aagctcctcctgagaaatatacaacacagatgtgagatgaagaagcagcagcttcttcg 1767  
3505 agacgttaccacacaaatataatgaacaaagatgtggaagataaaggagacatcgtccagag 3564  
1768 gccagaacaaagactagctattgtcaaggtccttctccaaaaccccaaatattatgt 1827

Db	3565	gtcaaaacaggagatgtgcatgtgccgagccctcatcagaacactcaacccctctgt	3624
Qy	1828	tgaatgagcgaactcaagccctcgataatgacagtgaaagtggtcaagcatcccttg	1887
Db	3625	tgtatgagcatcactcagctcttgatcactgaaagttaaagaattctccaagaagccctg	3684
Qy	1888	ataaagccagacgaggaagacatgctcgtgtgcaactcaagcgtctctgtcaattcga	1947
Db	3685	acaaagccagagaagcgcgacctgattgtgattgtctcacgcctgttccacatcaga	3744
Qy	1948	acggaagtttgatagtggtcttcgcaaaatggaaagaataaagaaacgaactatcaag	2007
Db	3745	atgagagcttaatactggtgttcttcagaaatgggagatcagaatgacagcatcagcc	3804
Qy	2008	agctctcgaaatcagacatatattttaagtttaagtgaatgcaagtcagt	2060
Db	3805	agctgcgcgacacgaagaagcatctatttcttcattcaatggtccagtgtccagcttgg	3857

## RESULT 2

ID AAZ888974 standard; DNA; 3924 BP.

AA AC AAZ88974;

DT 30-MAY-2000 (first entry)

DE Human MDR-3 DNA.

KM Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;  
KM stomach; colon; ulcerative colitis; pouchitis; large intestine; human;  
KM inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy  
KM ileum; anti-ulcer; anti-inflammatory; cystostatic; ss.

Homo sapiens.

AA WO200007577-A2.  
PN

PD 17-FEB-2000.

06-AUG-1999; 99WO-DE02426.

PR 06-AUG-1998; 98DE-1035526.

XX	1	1000	1000
XX	2	1000	1000
XX	3	1000	1000
XX	4	1000	1000
XX	5	1000	1000
XX	6	1000	1000
XX	7	1000	1000
XX	8	1000	1000
XX	9	1000	1000
XX	10	1000	1000
XX	11	1000	1000
XX	12	1000	1000
XX	13	1000	1000
XX	14	1000	1000
XX	15	1000	1000
XX	16	1000	1000
XX	17	1000	1000
XX	18	1000	1000
XX	19	1000	1000
XX	20	1000	1000
XX	21	1000	1000
XX	22	1000	1000
XX	23	1000	1000
XX	24	1000	1000
XX	25	1000	1000
XX	26	1000	1000
XX	27	1000	1000
XX	28	1000	1000
XX	29	1000	1000
XX	30	1000	1000
XX	31	1000	1000
XX	32	1000	1000
XX	33	1000	1000
XX	34	1000	1000
XX	35	1000	1000
XX	36	1000	1000
XX	37	1000	1000
XX	38	1000	1000
XX	39	1000	1000
XX	40	1000	1000
XX	41	1000	1000
XX	42	1000	1000
XX	43	1000	1000
XX	44	1000	1000
XX	45	1000	1000
XX	46	1000	1000
XX	47	1000	1000
XX	48	1000	1000
XX	49	1000	1000
XX	50	1000	1000
XX	51	1000	1000
XX	52	1000	1000
XX	53	1000	1000
XX	54	1000	1000
XX	55	1000	1000
XX	56	1000	1000
XX	57	1000	1000
XX	58	1000	1000
XX	59	1000	1000
XX	60	1000	1000
XX	61	1000	1000
XX	62	1000	1000
XX	63	1000	1000
XX	64	1000	1000
XX	65	1000	1000
XX	66	1000	1000
XX	67	1000	1000
XX	68	1000	1000
XX	69	1000	1000
XX	70	1000	1000
XX	71	1000	1000
XX	72	1000	1000
XX	73	1000	1000
XX	74	1000	1000
XX	75	1000	1000
XX	76	1000	1000
XX	77	1000	1000
XX	78	1000	1000
XX	79	1000	1000
XX	80	1000	1000
XX	81	1000	1000
XX	82	1000	1000
XX	83	1000	1000
XX	84	1000	1000
XX	85	1000	1000
XX	86	1000	1000
XX	87	1000	1000
XX	88	1000	1000
XX	89	1000	1000
XX	90	1000	1000
XX	91	1000	1000
XX	92	1000	1000
XX	93	1000	1000
XX	94	1000	1000
XX	95	1000	1000
XX	96	1000	1000
XX	97	1000	1000
XX	98	1000</	

XX

XX  
T T  
(M T 2010 T 20)

XX	3
XX	4
XX	5
XX	6
XX	7
XX	8
XX	9
XX	10
XX	11
XX	12
XX	13
XX	14
XX	15
XX	16
XX	17
XX	18
XX	19
XX	20
XX	21
XX	22
XX	23
XX	24
XX	25
XX	26
XX	27
XX	28
XX	29
XX	30
XX	31
XX	32
XX	33
XX	34
XX	35
XX	36
XX	37
XX	38
XX	39
XX	40
XX	41
XX	42
XX	43
XX	44
XX	45
XX	46
XX	47
XX	48
XX	49
XX	50
XX	51
XX	52
XX	53
XX	54
XX	55
XX	56
XX	57
XX	58
XX	59
XX	60
XX	61
XX	62
XX	63
XX	64
XX	65
XX	66
XX	67
XX	68
XX	69
XX	70
XX	71
XX	72
XX	73
XX	74
XX	75
XX	76
XX	77
XX	78
XX	79
XX	80
XX	81
XX	82
XX	83
XX	84
XX	85
XX	86
XX	87
XX	88
XX	89
XX	90
XX	91
XX	92
XX	93
XX	94
XX	95
XX	96
XX	97
XX	98
XX	99
XX	100

PT using phosphatidylcholine having mucous membrane protective activity to  
 PT prevent or treat large intestinal diseases, such as ulcerative colitis  
 PT or pouchitis -  
 XX  
 XS Disclosure, Page 17-22; 22pp; German.

CC This invention describes a novel/medicament containing a therapeutically  
CC effective amount of phosphatidylcholine to treat diseases, where  
CC phosphatidylcholine has an advantageous mucous membrane protective  
CC activity in the large intestine. The invention describes (1) a method  
CC to regulate presence or absence of MDR(multidrug resistance). 3  
CC transcription in a sample from patients with suspicion of stomach and  
CC colon diseases, where treatment with phosphatidylcholine is indicated;  
CC (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large  
CC intestinal inflammation, Crohn's disease, disease colitis, infectious  
CC enteritis/colitis, inflammation through X-ray treatment, antibiotics,  
CC chemotherapeutics, drugs or chemicals or large intestinal carcinoma,  
CC contains MDR3-specific primers to detect mutations in the MDR3 gene  
CC family and their interactions partners (in particular transcription  
CC factors); and (3) a method to manufacture a medicament for local gene  
CC therapy in the ileum with the ideal MDR3-analogous genes for patients

CC with missing or decreased expression of MDR3-analogous proteins. The  
CC products of the invention have anti-ulcer, anti-inflammatory and  
CC cyrostatic activity. The orally administered delayed-release form of  
CC phosphatidylcholine prevents premature resorption and provides for  
CC targeted release in the lower section of the small or large intestine.  
CC This sequence encodes the human MDR-3 protein described in the method of  
CC the invention.

Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Query Match	35.88;	Score 739.8;	DB 21;	Length 3924;
-------------	--------	--------------	--------	--------------

Matches 1268; Conservative 0; Mismatches 792; Indels 33; Gaps 2;

QY 1 cgagcaaggtcgcactacaatcgtgtgtagcacaccgacttctactatcgaagtgcag 60

Db	1765	ccaggaagcgccgacaccacatctgtgatacacacgacgtctctacgcttcgaaatgacg	1824
Qy	61	attgattgttacccttaagaagtatgaaatgtgttcgcggagaaaggaacattgtctgaatac	120
Db	1825	atgtcatcgtcgtgtcttgagagatgagtaattcttgagcagaaggaagccacagcgaactga	188
Qy	121	tcgcaaaacgggcctatattattcaacttggat-----	154
Db	1885	tgaagaagaagagggtgttactctcaactctgttcaacatgacagacatcaagaaagccagatcc	1944
Qy	155	-gtccagagatattaaagaagctgtatgaacagatggagttcaatlyacataattctactgaa	213
Db	1945	agtcagaagaatttgaactaaatgatgatgaagaagctgccactagaatgycgccaaatgct	200
Qy	214	gaagagccaactcactctctctgtcacctctgtgaaagcgatcaagtgtaagacttcatgaa	273
Db	2005	ggaactctgcctctatttagcatctctactcaagaanaaccttaanaaattcacaaatgtgtc	206
Qy	274	aggc-----tgaagatcccaaccacactcaagaagatgaagtctctcgtgaagctctcat	327
Db	2065	agaaagagccttgatgtgtgaaaaacgagatgacttgaagcaaatgtgccacagtgctcttc	2122
Qy	328	taaaaattttaaagttaaacaaagcctgaatgtgccttctgtgtctcgtggacatgtgctt	387
Db	2125	tgaaggtctctgaactgaataaaacagaatgycacctacttctgtcgtggagacagtatgtg	2184
Qy	388	ctgtctcaaatgagacgtttcatccagtatcttccaactcttctgnaaaattatacca	447
Db	2185	ccattgccaatgtgggtggtcttcaacggtcatttccagtcacattctcaagaatcatagcga	2244
Qy	448	tgtttggaataatgataaaccacatctaaagcatgatgtcagaagaatttatccatgatat	507
Db	2245	tttttgccacggagatgatgagcagtgagaagcagaagtgcaacatcatcttccttgattt	2304
Qy	508	tcgtcaatttgggtgttaattgtcttgcttcagttatctcatgcagggataatttlaagcga	567
Db	2305	tcttttctctgggaattattcttcttcttacttcttctctccacgggtttccacgtttbga	2364
Qy	568	gagcaggggaattttaagatgagatctaaacaacttggccttcaagccagctatatac	627
Db	2365	aagctgtgcgagatccccaacccaagagactgtgcgtcaatgtgtttaaagcaatgtcctaagc	2424
Qy	628	aggaatatgcctgttcttgatbaaaagaanaacagccagaaggtctgtgacaacaatatag	687
Db	2425	aggaacatgagcttgtttgatgacacataaaacagctactgtgtgaacttctacaagaactg	2488
Qy	668	ccatagatatagcaaaaattcaagaggcaacaggttccaggaatgtgcgtcttaacacaa	747
Db	2485	ccacagatgtgcaccaagttcccaagagccacgaacacaggtgtgcttcaattatgacaga	2544
Qy	748	atgcaactaataatgagccttccagtatcaattcctttatataatgatatggaggaagact	807
Db	2545	ataatagcaactcttggaacgtgttatatcatcatcatctactaaggttgcgcgttcaacc	2604
Qy	808	tcctgatctcgtatgtatgtctccagtaacttccgtgacagaagatgatgaaacccgacga	867

```

Db 2605 tattgtattatgacgtgtgttccaattatctgtgtcaggaattgttgaatgaattgt 2664
QY 868 tgaatgatttgcacaaagaataagaacacttaagcgtcgtgaagaatagaactgt 927
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2665 tggcttggaatgcacaaagaataaaaaagaactcggagcgtcgtggaagaattgtcaacag 2724
QY 928 aagcttggagaataatagctatagtgltcaatacagaaggaagaaagccttcagcaaa 987
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2725 aggaataagaataatattagacagctgtgtcttgcaccaggaagaatattgaatcaa 2784
QY 988 tttatgaagaatgtcttcaagctcaacacagaataactcgtgaagaacacagaattatg 1047
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2785 tttatgttgaataatgttatgtgaccttacaagaaattcgtgtcagaagaacacacattatg 2844
QY 1048 gaactgttatgacatgcacgtccatcttatatttgcctacagcagggtttcagat 1107
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2845 gaattacttattagatattccaaagaacttatgtatttcttctcagcgggttcttctgat 2904
QY 1108 ttggaagcatttaattcaagcttgaacgaatgacccacagaggaattgtcattttta 1167
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2905 ttgtgtcatatcatctgtgtatgtacatatgcgtcttcagagatgttatctgtgtttt 2964
QY 1168 ctgcaattgtcatatgagagcttgcacatcggaaaaagcgtctgttgcctcctgaatt 1227
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2965 ctgcaattgttatgtgtcagtggtccttagacatgccaagtccatttgcctcagactatg 3024
QY 1228 ccaagaccaaactcgggggtcgtgcacatcgttgcctgttgcgaagaagaaccaaataag 1287
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3025 ctaagcctaagcgttctgcagccacacttatcatcgtgttgcgaagaagaacactcgtatg 3084
QY 1288 acagccgcagtcacagaaggaagaaagccacagacatgtgaagggaatttagaatttcag 1347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3085 acagctacagtcgaagaggggtcgtgaagcctgaataattgaagaataatacatatttaag 3144
QY 1348 aagctcttcttcttatacctatcgtcccaagtggtttatctcctcgtgtgtatccctca 1407
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3145 aagctcgttcttaactatccaccggaacgacgtgcagtgcttccaggggtctgaagcctg 3204
QY 1408 gttatgagcggagaaagacagtagcatctgtgtggagcagcgtgtgtggagaagcactt 1467
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3205 aggtgaagaagaagccagacacatagccctgtgtggcagcagcgtgtgtggagaagcagg 3264
QY 1468 ctgtcaactctcgcagagactttatgaccccgctgcagaagcaagtgctgttgaatgtg 1527
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3265 tggctcagcgtcctgtgagcgtgttcaagcccttgcggggagacgtgtctcgtatggtc 3324
QY 1528 tggatgcaaaagaattgaattagatgagtggtccgttcccaatagcaatcgttctcctcaag 1587
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3325 aagaagcaaaagaactcaatgttccagtgtcagagctcaactcggaaactgtgtctcag 3384
QY 1588 agcctgtgctcttcaactgcagactgtgtcgaacatgcctatgtgtgaacaacgcgtg 1647
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3385 agctatcctatttgcctgcagcgtcgtcgtgagatatttgcctatggaagcaacacgcgg 3444
QY 1648 tgggtgcaattagatgagatcaagaagccgcaaatgtcagcaaatatccattcttattg 1707
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3445 ttgtatcacagagatgaattgtgagtgcaagcacaagctgcacaatatatccttcatcg 3504
QY 1708 aagcttccctcgtgaagaataacacacaagaattgtgactgtgaagggcagcagcttctgcg 1767
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3505 agacgttaccacacaataataaagaagtgagataaagggaactcgaagcttccagagag 3564
QY 1768 gccggaacaaagaactagctattgtcaagggctcttctcacaacaccccaatttatgt 1827
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3565 gtcaaaaacagagatgtgctattgtccgagccctcaltacagcaactcctcctcgt 3624
QY 1828 tggatgagggcacttcagccctcgataatgacagtgagaaggtgtgtcagatgcctctg 1887
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3625 tggatgagagctacatcagctctgtactgaagtgaaaggtgttccaagaagccctcgg 3684
QY 1888 ataaagccagagcgggaagagacatgctagtgtcactcagcagcgtctctgtcaattcaga 1947
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3685 acaaaagccagagagcgcacctgtcattgtgtcaccgctgtccacacatccaga 3744

```

```

QY 1948 acgcagattgatagtggttctgcacaaatggaaagataaagaagaactcatcaag 2007
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3745 atgcagacttaattagtggtgttccgaattggagagtcagaaggaatgagcagcgtcagc 3804
QY 2008 agctcctgaagaatcgcagacataatttaagttagtgaaatgcagtcagtg 2060
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3805 agctgtcgcacagaagaagcactatatttcaatcagtgctcagtgctcagcgtcgg 3857

RESULT 3
AAQ52726
ID AAQ52726 standard; cDNA; 4669 BP.
XX
AC AAQ52726;
XX
DT 24-JUN-1994 (first entry)
XX
DE Sequence of human multi-drug resistant gene mdr1.
XX
KM P-glycoprotein; multi-drug resistant gene; mdr1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 424..4267
FT FT /*tag= a
XX
PN W09324613-A.
XX
PD 09-DEC-1993.
XX
PE 14-MAY-1993; 93WO-US04707.
XX
PR 22-MAY-1992; 92US-0887712.
XX
PA (GENE-) GENETIC THERAPY INC.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI McDonagh KT, Nienhuis A, Tolstoshev P;
XX
DR WP1; 1993-405805/50.
DR P-PSDB; AAR44297.
XX
PT DNA or RNA sequence for human multi-drug resistant gene MDR1 -
PT encodes P-glycoprotein with altered splice site, used in cancer
PT therapy
XX
PS Example; Fig 4; 64pp; English.
XX
CC PMDR2000 contains an mdr1 cDNA sequence (AAQ52726) described in PCR
CC application no. W087/0943m wherein the first 282 bp of the 5' UTR
CC and the last 23 bp of the 3' UTR of the cDNA sequence have been
CC removed. The codon YTN is unclear in the spec., and encodes the AA
CC Leu.
XX
SO Sequence 4669 BP; 1394 A; 893 C; 1126 G; 1254 T; 2 other;

Query Match 34.5%; Score 712.6; DB 14; Length 4669;
Best Local Similarity 60.0%; Pred. No. 1.8e-181;
Matches 1261; Conservative 1; Mismatches 796; Indels 42; Gaps 3;
QY 1 cgaagcaaggtcggactacaatcgtgtgtagacacggaacttctactatcgaagtgcag 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2151 ccagaaaaggtcggaccacccttgtgatagtcactcgttctcagcgttctgaatcgtg 2210
QY 61 attgatgtgaccctaaagaatgaaatgctgcggcggagaaagagacatgtaactaa 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2211 acgtcatcgcgtgttctgatatgagtgatcattgtcgggaaaggaatcatgataactca 2270
QY 121 tggcaaaacgaggtcatattatcacttgcattgtcagtgctacagagatataaaagctgat 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 2271 tgaagagagaagcatttacttcaaatctgtlccaaatgcaagcagcaggaattgattg 2330  
QY 181 aacagatgagtcataatgacatatctta-----ctgaagaagaaccaa 223  
Db 2331 aattgaanaattgcagctgatatccaaaagtgaattgatgcttggaaatgcttcaa 2390  
QY 224 ctccattctctgcgcctctgtgaagagcatcaag-----tcag 261  
Db 2381 atgatttccaaatccagcttcaataaagaataaagatcaactcgttagagtgctcgtgatcc 2450  
QY 262 acttcattgacaagcgtgagaaatccaccacatctaaagagataagcttccctgtagtct 321  
Db 2451 aagcccaagacagaagctttagtaccaaagagcctcgtgatgaagataactccagttt 2510  
QY 322 ctctattaaaaatttaagttaaaccaagcctgaatgagccttctgtggtctggggaat 381  
Db 2511 ccttttgagagatttagaagcaaaattcaactgaatgagccttatttgtgtgtat 2570  
QY 382 tggctctgttctaattgaagctgttccatccagtatcttccatcttcttgaanaatta 441  
Db 2571 ttgtgccaattataaattggaagcctgcgaacacagcatttgcataaatttccaagaatta 2630  
QY 442 taaccattgttgaataataatgataaaccacattaaag---catgatgcagaatttatt 498  
Db 2631 taggggttttcaagaagattgtgatccctgaacaaagagacagataagtaactgttct 2690  
QY 499 ccaatgatttcgtcatttgggtgttatttgccttgcagttattcattcagcagattat 558  
Db 2691 caactatgttctcagcccttggaaattatttcttatttattcatttccctcaggtttca 2750  
QY 559 tttaaggagagcagagggaattttaagatagatttaagacacttggcttcaagaacca 618  
Db 2751 caattgccaagaagctggagagatccccaacaaagcgtcccgataaagatttccgatcca 2810  
QY 619 tgtlatcagaagatactgcctgtgttgaatgaanaagaaacagacagaggcttgcaca 678  
Db 2811 tgcctagacagagattgtgattgtttagtacccttaaaaaacacacttgtagatgcata 2870  
QY 679 caattttagcagataatagcaaaatccaagagcaagcaggttccagattgagcttct 738  
Db 2871 ccaggctcgcacatgtctgtcgaagttaagggctataaggttccaggtctgctgtaa 2930  
QY 739 taaccaaaatgcacactaacatgagacttccagttacattccattccattatatttgg 798  
Db 2931 ttaccagaataatgaacaaacttgggaacaggaataattatattccattcacttgcctg 2990  
QY 799 agatgacattccctgattctcgtagttatgctccagttactgcgtgacagaagattgaa 858  
Db 2991 aacttaacactgttactcttagcaattgttaccatcatcttgcataagcagagttgttga 3050  
QY 859 ccgcagcaatgacctggaatttgcacaagaataagcaaaacttaagcattgctggaaga 918  
Db 3051 tgaanaatgtctgtcagacagacactgaagaataagaaactagaaggtgctvggaaga 3110  
QY 919 tagcaactgaaccttggagaatatagctactatagtgcttataaanaaggaaaagcct 978  
Db 3111 tgcctactgaagcaaatagaacacttccgaaccgttgttcttgcattcagagcagaagt 3170  
QY 979 tcgagcaaatgtatgaagaagatgcttcaactccaacacagaataactcgaagaagcac 1038  
Db 3171 tgaacaatatgtatgtctcagagtttgcagtgatccatacagaacacttggagaaagcac 3230  
QY 1039 agattattggaagctgttatgcaattcagcattgcttataatttgcctatgcagcag 1098  
Db 3231 acactcttgaattacatttcccttccaccagcaatgatttatttccatagctgagat 3290  
QY 1099 ggttcgatttggagcctttaaattcaagctgagcagatgagcccaaggagcatttca 1158  
Db 3291 gttccggatttggagcctactcgtgtgacataaaatunatgagccttggagatgtctgt 3350  
QY 1159 tagtttactacgaattgataatgagctatgccaatcgaaaacagctcgttcttgctc 1218  
Db 3351 tagtatttccagctgtgtcttcttggcagatgcccgtggggaagcagttcaatttgcctc 3410

QY 1219 ctgaatttccaagaagccaatcgggggctgcgcatctgttgccttcttggaaaagaaac 1278  
Db 3411 ctgactatgccaagaagccaataataatcagcagcccaatcatcatgattgaaaacccc 3470  
QY 1279 caaataagacagccgaotlcaagaagggaaaagccagacacatgctgaaggaatttag 1338  
Db 3471 ctttattgacagcttaccagcagcaggaagccttaatgctgcgaacacatttggaaatgtca 3530  
QY 1339 agttcgaagaatctcttcttctatccatgctgcgccagaatgttccatccctcgtgct 1398  
Db 3531 catttggtaagtattatccaattccaccgcgcagacatcccaatgcttccagggac 3590  
QY 1399 taccctcagtttagcgaagaaagacagatgcatcttgttggggagcagcgtcttggga 1458  
Db 3591 tgaagcttggaggtgaagaaggccagagcctggtccttgggtggcagagtggtctggga 3650  
QY 1459 aagacattcttccaactctgcagagactttagaccctgtgcaagagacatgctgt 1518  
Db 3651 agagcagatgtgtccagctcctcgtgagcggttctacgaccccttgcagaggaagtgctgc 3710  
QY 1519 ttgatgtgtgagtgcaaaagaattgaatgtacagtgctccgttcccaatagcaatcg 1578  
Db 3711 ttgatgcaagaataaagaacgcatgaattgtcagtggtcctgagacacactggtgcacg 3770  
QY 1579 ttccccaagagcctgtgcttccactcagctgagcattgtgtgagacatcgctatgtgaca 1638  
Db 3771 ttccccaagagcctcacttcttgcattgagcattgtgtgagacatcgctatgtgaca 3830  
QY 3771 ttccccaagagcctcacttcttgcattgagcattgtgtgagacatcgctatgtgaca 3830  
QY 1639 acagccgtgtgtccatagattagatcaagaagccgcaaatcagcaaatatccatt 1698  
Db 3831 acagccgggtgtgtgtccacaggaagatcgttgaaggcagcaagaaggaacataatag 3890  
QY 1699 ctttattgaagcttccctccttgaataatacaacacacaaagtgtgactgaagaagcagac 1758  
Db 3891 ccttcaatcagatcagctgcttaataatagacataaagtagagacaaaggaactcagc 3950  
QY 1759 ttcttgcgcgcagaaacaaagactagctatttgcagaagccttcttccaaaacccaaa 1818  
Db 3951 tctctgtgtgcagaaacaaacgcatgttgcctagctgtgccttgtttagacagccctcata 4010  
QY 1819 tttaattgttgaatgagccacttccagccctcgataatagctgaaggtgtgtcagc 1878  
Db 4011 ttgtgctttagatgaagcagcagctcagctctgtgatacagaaggtgaaggtgtccag 4070  
QY 1879 atgcccctgataaagccagagcggaagaacatgcttagtgcctacacagccttctc 1938  
Db 4071 aagccctggaacaaagcagagagcgccagcctgtgtgtattgttgcacgcctgttcca 4130  
QY 1939 caattcagaagcagattttagatagtggtcttgcacaaatggaagataaagggaacagga 1998  
Db 4131 ccatcagaatgtcagactttaaattagtggttccaagaatgagagcagaagagatgga 4190  
QY 1999 ctcatcaagagctcctgagaaatcgagacataattttagtaagtgaatgacagctcag 2058  
Db 4191 cgcattcagcagctgctgtgcacagaaagcattcatttccaatgtgtcagtgctcagctg 4250

RESULT 4  
AAQ70907  
ID AAQ70907 standard; DNA; 2726 BP.  
XX  
XX AAQ70907;  
AC  
XX  
XX  
DT 27-MAR-1995 (first entry)  
XX  
XX  
DE Multidrug-resistance gene MDR-1.  
XX  
XX  
KW Multidrug-resistance; MDR-1 gene; ds.  
XX  
XX  
OS Synthetic.  
PN W09417206-A.



```
Db 1911 acagccgggtggtgtcacagggaagatcgtgaggcgacgcaaaaggccaacatactg 1970
Qy 1699 cttattgaaggtctccctgagaataacaacaagttgactgaaggagcacagc 1758
Db 1971 ccttcacgagtcactgctataataatagacactaaagtagagacaagaactcagc 2030
Qy 1759 ttcttgaggcgacgaanaaagactagctattgcaaggtccttccaaaacccaata 1818
Db 2031 tctctgggcccgaataaacaagcattgccaatagctcgccctgtgttagacgcctcata 2090
Qy 1819 ttctattgtgatgagccacttcagccctcagataagacagtgagaagtggttcagc 1878
Db 2091 ttgtgcttttgatgagaacccagctcagctctgatacagaagaagtgaagttgtccag 2150
Qy 1879 atgcccctgataaagccaggaacgaggacatgctcctagtgtcactcaacagctctctg 1938
Db 2151 aagccctgacacaagccagagaagcgccagcctgcatgtgtgtcaccgcctgtcca 2210
Qy 1939 caattcagaagcgagatttgtagtgttctgcacaatggaagaagtaagaagcaaggaa 1998
Db 2211 ccatcagaatgacgaacttaataagtggtctcagaatcgcgagtcagaagggatggca 2270
Qy 1999 ctcatcaagagctccctgagaaatcgaacataataatttaagttagtgaatgcacagtcag 2058
Db 2271 cgcacacagcagctgctgcacacagaagaagcatttcttaaatgtgtcagtgctcagctg 2330

RESULT 5
AAO70916
ID AAO70916 standard; DNA: 2726 BP.
XX
AC AAO70916:
XX
DT 27-MAR-1995 (first entry)
XX
DE Multidrug-resistance gene MDR-1.
XX
KM Multidrug-resistance: MDR-1 gene; ds.
XX
OS Synthetic.
XX
PN MO9417210-A.
XX
PD 04-AUG-1994.
XX
PE 25-JAN-1994; 94MO-US00748.
XX
PR 27-JAN-1993; 93US-0010433.
XX
PR 16-DEC-1993; 93US-0168621.
XX
PA (ONCO-) ONCOR INC.
XX
PI Bhattachar SK, George AL;
XX
DR WPI: 1994-264122/32.
XX
PT Enzymatic amplification of target nucleic acid sequences to form
PT exact or modified copies - has increased fidelity and can identify
PT point mutations or allele(s)
XX
PS Disclosure: Page 30; 69pp; English.
XX
CC This gene is used as an example of a target DNA in a method for the
CC amplification of nucleic acid sequences in a mixture using various
CC DNA probes/primers. The method may be used to detect a mutation or
CC allele in the target, e.g. that that causes a genetic disease.
XX
SQ Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;
```

```
Qy 1 cgagcaaaaggttcgagctacacatcgtgtgtagcacaccgactttctactatttgaagtgcag 60
Db 231 ccagaaaaggttcgagccaccatctgtatagctacatcgtttgtctacagttgttaattgctg 290
Qy 61 attgattgtgacccctaaaggaatgaaatgctgtggtgggaaagagacacatgtgactca 120
Db 291 acgtatcgtcgtgttcgtatgtatgtgagtcattgtgtggaagaagaatacatcatgtgaccca 350
Qy 121 tggcaaaaagagttatattatattcactgtgtatgtcacaagatatataaagctgtgac 180
Db 351 tgaagagagaagagcatttaacttcaaacctgttcacaaatgcagacagcaagaatgaaatgac 410
Qy 161 aacagatgagatcgaatgacattctca-----ctgaagaagaaccaa 223
Db 411 aattgaaaatgtcagctgatgtgaatccaaaagtgaattgtatgctcttggaaatgtgtccaa 470
Qy 224 ctcaactcctctgcactctgtgaagagcaatcaag-----tcag 261
Db 471 atgattcaagaatccagttcctaataaagaaagatacactcgtatgaggtgtccgtgtgatac 530
Qy 262 acttcatgacaaagctgaggaatccaccacaatcctaaagagataagttcctcgaagttc 321
Db 531 aagcccaagacagaagaagcttagtaccacaagaagctcgtgatgaagtataccctcagttt 590
Qy 322 ctctattaaaaatttaaaatgaataaacagcctgaatgaccttgtgtgtcgtggagact 381
Db 591 ccttttgagagattatgaaatgaatattcaatgaaatgacctatttctgtgtgtat 650
Qy 382 tggctctgtcttaaatgaaactgtgtcatccagatatttccatcactcttgaanaatha 441
Db 651 ttgtgtcattataatgaaagcctgcgaacacgaatgtgcaataatatttcaaaagatta 710
Qy 442 taacattgtttgaaataatgataaacaacataaag---catgatgcagaatattat 498
Db 711 taggggttttacaagaattgtatgtatcctgcgaataaacaagaacagatagtaactgtttt 770
Qy 499 ccataatattcgtcatatttgggtgtatattgtctgtgcagtaatttcatgcaggaattat 558
Db 771 cactattgttttaagcccttgaattatcttcttatttcaatttcttccctcagggtttca 830
Qy 559 tttaagcagaagcagaaggaattttaacgataagatttaagacactgtgacctgcaagcca 618
Db 831 cacttgcagaagctgagagatccccaacagcggctccgatacatggtttccgatcca 890
Qy 619 tgtatcatcagaatattcgtgtgttgaatgaagaagaaacagcagaaggtctgacaa 678
Db 891 tgcacagaagaagtgtgagttgtgttgaatgacccaataaacaacactgtagcattgaccta 950
Qy 679 caattatgacataagataagacaacaattcaagagcaacaggttccagatgtgcgtct 738
Db 951 ccaggtcgcgcaatgtgtcctcaagttaaaagggtctatagttccaggtcgtgttaa 1010
Qy 739 taacacaaatgcaactaactaaggtgacttgcattatcattccttataatgtatgg 798
Db 1011 ttaccagaataatagaacaactcttggaacagaagaataatataccttcaactatgtgtg 1070
Qy 799 agatgacattcctgtattcctgagatgtgctcagtaacttccgtgcgaagaatgatgaa 858
Db 1071 aactaactgttactccttagcaattgttaccatcatcttgcaatagcagagttgttga 1130
Qy 859 ccgacgaatgactgatttgcacaacaaagataagcaagaacttaagcactgtgaaga 918
Db 1131 tgaaatgtgtcgtgacacagcactgaagaataaagaagataaggtgtgtggaag 1190
Qy 919 tagcaactgaagcttttggaatatacgtactatgtgtcattacaaggggaaaaagcct 978
Db 1191 tgcgtactgaaagcaatagaacaacttccgaacggtgttcttgcatacgaagcgcaag 1250
Qy 979 tcgagcaaatgatatgaagatgcttccagactcaacacagaataactcctgcagaagaagc 1038
Db 1251 ttgaacataatgtatgtcagagtttgcagagtacacataagaaactcttggagaaagc 1310
```



[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 1262; Conservative	60.1%;	712.4;	18;	2726;
		Pred. No. 1.6e-181;		
		Mismatches 796;	Indels 42;	Gaps 3
1	cgagcaaaagtcgactacatcgtggtgacacacccgacttctactatctgaagcgag	60		
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				

[illegible]

QY	1219	ctgtaattccaaagccaaatcggggtctgcgcatctgttttgcttctgtttgtaaaagaac	1278
Db	1491	ctgactcttgcgaagccaaataatcagcgcccaactcactatgatacttgaataaaacc	1550
QY	1279	caaatatagaacggccgagcacaagaagggaagaaagccaaacatgtatgaaggaattag	1338
Db	1551	ctttgatctgacagctaacagacggaagcgcttaatgcggaacacatgtgaagaaatgtca	1610
QY	1339	agtttcgagaagctctcttctctatccatgtctgcgccgaatgttttcaatctccgttgct	1398
Db	1611	catttggctgaagtgttacttaactatccaccgcagccagacatcccaatgtcttcaggac	1670
QY	1399	tatccctcagatattgagcgagggaaagacagtacatttgggggacgcggtgtggga	1458
Db	1671	tgaagcttgaaggtgaagaagggccaaacgcgtctctctgtgtggcgcaagctgtgcgtgga	1730
QY	1459	aaagcaactctctgtcaactctctgcagaagacttatagaccggctgcagaagacaagctgt	1518
Db	1731	agagcacaggtgtcccgctctctctgtgagctgtctctacgacccctctgcagggaaagtctgc	1790
QY	1519	ttgatgtgtgtgagtgcaaaagaattgaaatgtacagtgctccggttcccaatagcaatcg	1578
Db	1791	tgtatgtgcgaagaataaagaacgacatgaattgaattaggtgtcctcgagcacaccctgtgc	1850
QY	1579	ttctccaaagggccgtgtcttcttcaactctgacgaattgtcttgagagaacatgcctatgttaca	1638
Db	1851	tgcccacaggaecccaaccctctgttactctgcagacatgtctgtgagaaacatgtccctatgtgagaca	1910
QY	1639	acagacgtgtgtgtcccatctagatgtagatctcaaaagaagccccaatgtcaagcaaatatccat	1698
Db	1911	acagcggtgtgtgtctcacaggaagatcgctgtgaggtgcagcaagaaggccaaactatacgt	1970
QY	1699	ctttatttgaaagtctctccctgagaataatacaacacagtttgactgtgaagagagcaagc	1758
Db	1971	cccttcatctgagctcgtccctcaataatataatagcactaaatgaagagacaagaagaaactcagc	2030
QY	1759	ttctctgcgagccagaacaagaactagctcatctgtgcgaaggtctcttcccaaaacccaana	1818
Db	2031	ctctctgtgtgcgcagaaacaacagcatgtgcataagctctgcctctgttagacagcctcata	2090
QY	1819	ttttattgtgtgatatgagccacttccagccctcogataatagaacgtgagaaggtgtgtcaagc	1878
Db	2091	ttttgtctttttgatgtagaagccagctcgtctgtgatacagaagaagtgaanaaggtgttccaa	2150
QY	1879	atgccccttgataaagcgcaagcagggagaagacatagctcttagtgttcaactcaagcgtctct	1938
Db	2151	aagccctgtgaaacaagccagaggaagccgcagccgcagctgtgtgtcttcaaccgctgttcca	2210
QY	1939	caatttcgaagcgagattttatagtgttctctgcacatgtgaagaagataaaggaacaagaa	1998
Db	2211	ccatccagaatgcagacttaataatagtgtgtttcagaatgtgcagatgtcaagaagatgtgca	2270
QY	1999	ctcatcaagagctcttcyagaatactcgagacatatatttaagttagtgaatgcacagttcag	2058
Db	2271	cgatccagcagctctgtgcacagaagaagcattcatatttttcaatgtgtcagttgtccagctgt	2330

RESULT	7
AAH57442	
ID	AAH57442 standard; cDNA; 4349 BP.
XX	
AC	AAH57442;
XX	
DT	10-SEP-2001 (first entry)
XX	
DE	Human intestine cell specific cDNA sequence SEQ ID NO:282.
XX	
KW	Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW	lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW	metabolic disease; developmental disease; cytostatic; immunomodulatory
KW	neuroprotective; gene therapy; cancer; immunopathology; neuropathology
XX	
OS	Homo sapiens.

XX XX  
PN WO200132927-A2.  
XX 10-MAY-2001.  
PD  
XX 02-NOV-2000; 2000WO-US30396.  
PE  
XX 04-NOV-1999; 99US-0163508.  
PR  
XX (INCY-) INCYTE GENOMICS INC.  
PA  
XX Sornasse T, Sellhammer JJ, Watson GA;  
PI  
XX WPI; 2001-291057/30.  
DR  
XX  
PT New cell and tissue specific polynucleotides useful for diagnosis,  
PT prognosis or monitoring of treatments for disorders where the gene is  
PT associated with a cancer, immunopathology or neuropathology -  
XX  
XX Claim 1; Page 207-208; 327Pp; English.

XX  
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide  
CC sequences (I). (I) can have cytosolic, immunomodulatory and  
CC neuroprotective activities, and can be used in gene therapy. (I) and  
CC proteins (II) encoded by them are used in high throughput screening  
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,  
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or  
CC their fragments, immunoglobulins, inhibitors, drug compounds and  
CC pharmaceutical agents. Expression of (I) in a sample indicates the  
CC differentiation of embryonic stem cells into a tissue selected from  
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic  
CC tissues. (I) and (II) are used to produce an expression profile that  
CC defines a metabolic or developmental process, treatment, condition,  
CC disease or disorder. The gene profile can be used for diagnosis,  
CC prognosis or monitoring of treatments and for investigating a  
CC predisposition to a disorder where the gene is associated with a  
CC cancer, immunopathology or neuropathology.  
XX  
SQ Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Query Match 34.5%; Score 712.4; DB 22; Length 4349;  
Best Local Similarity 60.1%; Pred. No. 1.9e-181;  
Matches 126; Conservative 0; Mismatches 796; Indels 42; Gaps 3;

QY 1 cgagcaaaagtcgagcactcaatcgtgtagcagccagccttctacttctgaagtgcag 60  
DB 1852 ccagaaaagtcgagcagccatctgtgatagcctcgttctgctacagttcgtaatgctg 1911  
QY 61 attgattgtgaccctaaagatgtaabctgctgagcaaggaagacatgctgaacaa 120  
DB 1912 acgtcatcgtcgttctgtagatgagtcattgtgaggaagaaatcatgatgaccca 1971  
QY 121 tggcaaaaagcgtctatataatcacttctgtagtgcacagagatatataaagctgag 180  
DB 1972 tgaagagaagaaggaattactcaaatctgcacatgcagacagcaggaatgaagttg 2031  
QY 181 aacagatgagtcacatgacatctcta-----ctgaagaagaagaccaa 223  
DB 2032 aatagaagaatgcgcgtgataatccaaagtgaattgagccttggaaatgcttcaa 2091  
QY 224 ctcaactctctgactctgtgaagagcatcaag-----tcaag 261  
DB 2092 atgattcaagtcagcttaataagaagaatcactcgttagagtgctccgtgatacac 2151  
QY 262 actcatatgacaagcgtgaggaatccaccacatctaaagataagttcttcgaagctc 321  
DB 2152 aagcccaagaagaaggttgtaaccaagaagcgtctgtgataagataatcctccagttt 2211  
QY 322 ctctatataaattttaagttaacaagcctgaatgagccttctgtgttctgggaat 381  
DB 2212 ccttttgagattatgaagccttaactgaatgagccttattctgttctgtgatat 2271

QY 382 tgcctctcttaaatgaagactgttcatccagatatttccatctcatcttgcgaataatla 441  
DB 2272 ttgtgcattatataatgagagcctgcacaccagcatctgtcaataatatttccaagtta 2331  
QY 442 taaccatgttggaaataatgataaaacacattaaag---catgagcagaattattc 498  
DB 2332 taagggttttacaagaattgatcctctgaacaaacagacagaatagtaactgtttt 2391  
QY 499 ccatgatatctgcatttctgtgttattctgtctgtgcagttatctgaagagattat 558  
DB 2392 cactatgttctgaccttggaattattcttattatcaacttcttccatcaaggtttca 2451  
QY 559 ttacgcagagcaggggaaatttaacgatagataagaacacttgccttcaagcca 618  
DB 2452 catttggcaagcgtgagagatccctcaccagcgtctccatgattgttccgatcca 2511  
QY 619 tgttatatcagatattgctcgtgtttagtgaagaagaaacagacagagagcttgcaca 678  
DB 2512 tgcacagacaggaatgtgtgtttagtgaacccaacacacacttgcagcatgtgacta 2571  
QY 679 caatatagcatalagatagcacaaattcaagagagacaggttccagatctgcctct 738  
DB 2572 ccaggtccgcgaatgagtcgtctcaagttaaaagggctataggttccaggtctgttaa 2631  
QY 739 taacacaaatgcaactacaatgggacttcaatcatcttccattatataatgagatgg 798  
DB 2632 ttaccagaatagacaatcttgtagacagaataatataatcttccatcatgtgttgc 2691  
QY 799 agatgaccttccatattctgtagtattgctccagacttgcgcgtgacaggaatgttga 858  
DB 2692 aactaacacttactcttaacatctgtaacccatcatctgcaatgacaggaatgttga 2751  
QY 859 ccgcagcaatgagcaggaatttgcacaaagataagcaagaactaaagatctggaaga 918  
DB 2752 tgaanaatgttctcgtggaagaagcactgtaagaagataagaagactgaagtgctgga 2811  
QY 919 tagcaactgaagccttggagaataacgtactatagtgatgatacaaggaagaaagcct 978  
DB 2812 tgcctactggaagcaatagaaacacttcgaaccgtgttcttctgactcaagagaagat 2871  
QY 979 tcgagcaaatgataagaagatgcttccagactcaacacagaataatccctcgagaagaag 1038  
DB 2872 ttgaacatatgtatgctagaagtttgcaggtacacataagaagaactcttggagaagac 2931  
QY 1039 agattatggaagcgtgtatgcattcaacagcctgcttataatttccatcagaag 1098  
DB 2932 acactcttggaataacattcccttccaccagcaagaatgattatttcccatctgagat 2991  
QY 1099 gtttcgaattggagcctatttaattcaagctggaagaatgacccagagggcatgtca 1158  
DB 2992 gttccggttggagcctactgtgagcacaataacatcatgagcttgaagatgttctgt 3051  
QY 1159 tagttttacgcgaattgcatatgagatgagtcagcatctggaaaagcgtcttggctc 1218  
DB 3052 tagatttttcagctgtgtctgtgtgctgacggtggtggtggaagtcagttcatttgc 3111  
QY 1219 ctgaatatccaagccaactcgggggtgctgcactctgttgccttgttggaaagaagaa 1278  
DB 3112 ctgactatgccaagaacaaatalcagcagccacatcatcatgattcatatgaaaaaacc 3171  
QY 1279 caaatatagacagccgcatgtaagaagggaagaaagccagacacatgtgaaggaaattag 1338  
DB 3172 ctgtgatgacagctacgcacggaagcccaatgacgaacacattggaagaatgtca 3231  
QY 1339 agtttgaagaagctcttcttcatcatcattgtgcccagatggtttcatctccgtgt 1398  
DB 3232 catgtgtgaagttgtatctacacttccaccagccagcaggaatccagtgcttcaagac 3291  
QY 1399 tatccctcagatgagcaggaagagacagatgagcattgtgttgcagcgtcgttggaa 1458  
DB 3292 tgaagcctggaagtgaaagaagggccagagcgtgctgtgtgtgagcagcagtgctgtg 3351  
QY 1459 aaagcacttctgtcaactctgcagagacttlatgaccccggtgcaaggaagatgctgt 1518

```

Db 3352 agagacagctgttccagctccctggagcggttctacgaccttggcagggaagcgctgc 3411
Oy 1519 ttgatgctgtgatgcacaaagaatgatgatcagtgctccgtcccaataatgcac 1578
Db 3412 ttgatgcagaagaataaagcagctgaatgttcatgtgctccgagcacacccgggctgc 3471
Oy 1579 ttcccaagagcctgtgtcttccacttcagctgctgagaaacgtgcctatgtgtgc 1638
Db 3472 tctccagagagcccaacctgttctgactcgcagcatgtcgaagaacatltgcctatggagaca 3531
Oy 1639 acagcgcgtgtgctgcatatagatgcacaaagaagccgcaaatcatcaaatccatt 1698
Db 3532 acagcgcgggtgtgtgtcacaagaagagatcgtgagggcagcaagaagagcccaatactatg 3591
Oy 1699 ctttatatgaagctcctcctgagaataacaacacaaagtttgactgaaagagcagc 1758
Db 3592 ccttcacgcagctcactgcttaataatatagcactaaagtagagagacaagaagcactcagc 3651
Oy 1759 ttctcggcggcagcaaaacaaagactatgtcgaaggtccttcccaaaaacccaaa 1818
Db 3652 tctctggcggcagcaaaacaaagcalttgccatagctcgtgccttcttagacagcctcata 3711
Oy 1819 ttctatgtgtgatgagcaccctcagccctgcataatgcagtgagagtggtctcagc 1878
Db 3712 ttctgcttggatgtgaagccacgctcagctcgtgatcacagaagtgtaaaagttgtccaa 3771
Oy 1879 atgccttgataaagccagagcagcagacatgcctagtgtgtcactcagcagctctctg 1938
Db 3772 aagccctgcagcaaaagcagagagcgcagcactgtgctgtgtcctcagccgctgtcca 3831
Oy 1939 caattcagaacgcagatttgatagtggttctgcacaaatggaagtaaggaacaaagaa 1998
Db 3832 ccattcagaatgacacttaataagtggttctcagaatggtcagagtcagagagcatgtgaa 3891
Oy 1999 ctccacaaagcctcctggaataatgcagacatatatttaagtagtgaatgcagctcag 2058
Db 3892 cgcacacagcgtcgtgtgcagacagaagacatctatttctcaatggtcaggtgtccagctg 3951

RESULT 8
AAO72872
ID AAO72872 standard; cDNA: 4646 BP.
AC AAO72872;
XX
XX 21-JUN-1995 (first entry)
DT
XX
XX Human multidrug resistance gene (MRD-1) encoding a p-glycoprotein.
DE
XX
XX Human; tyrosinase; compound; hair; follicle; liposome; lipophilic;
KW lipophilic; stratum corneum; melanin; hair dye; glycoprotein; alopecia;
KW chemotherapeutic; transforming growth factor; growth stimulant; aromatase;
KW cyclosporin A; macromolecule; polymer; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 425..4267
FT /*tag= a
FT /product= p-glycoprotein
XX
XX WO9422468-A.
XX
XX 13-OCT-1994.
XX
XX 01-APR-1994; 94WO-US03634.
XX
XX 02-APR-1993; 93US-0041553.
XX
XX 13-JAN-1994; 94US-0181471.
XX
XX (ANTI-) ANTICANCER INC.

```

```

PI Li L, Lishko VK;
XX
XX WPI: 1994-332816/41.
DR
DR P-PSDB: AAO72872.
XX
XX Liposomes for delivering protein, nucleic acid etc. to hair
PI follicles - e.g. to restore hair colour, prevent hair loss during
PT chemotherapy, stimulate hair growth etc.
XX
XX Claim 8; Page 70-77; 100pp; English.
PS
XX
XX The nucleotide sequence of the human multidrug resistance (MDR-1) gene
XX which encodes a p-glycoprotein. This is an example of a compound which
CC can be delivered to hair follicles via a novel liposome composition.
CC The liposomes are comprised of a lipophilic or lipophobic compound which
CC will selectively target the hair follicle (via the stratum corneum)
CC without damaging or unwanted effects on cells outside the follicle.
CC Compounds e.g. p-glycoprotein, can be delivered to the hair follicle for
CC treatment of chemotherapy-induced alopecia. Other compounds targeted at
CC hair follicles can include: tyrosinase (or the DNA encoding it -
CC AAO72871), melanin or hair dyes (to restore hair colour or condition);
CC human transforming growth factor alpha (AAO72873) (for reversal of wavy
CC hair); hair growth stimulants (e.g. cyclosporin A or aromatase) or
CC antisense sequences. The method allows compounds (e.g. macromolecules
CC or polymers), which would not normally reach the hair follicles, to be
CC delivered to these target areas.
XX
XX Sequence 4646 BP; 1371 A; 893 C; 1128 G; 1254 T; 0 other:
SQ
XX
XX Query Match 34.5%; Score 712.4; DB 15; Length 4646;
XX Best Local Similarity 60.1%; Pred. No. 2e-181;
XX Matches 1262; Conservative 0; Mismatches 796; Indels 42; Gaps 3;
XX
Oy 1 cgagcaagtgctgactcaatcgtgtgtagcacaccgacttctactatcgaatgcag 60
Db 2151 ccagaaagagtcggcaccacattgtgatagcctcactgttgcacagtgctgaatgcgt 2210
Oy 61 atttgatgtgacccctaaagatgtgaatgctggtggagagaagagacatgctgaactaa 120
Db 2211 acgtatcgtgttctgctgtagtgatgcatgtgtgagaaagaatcatgtagactca 2270
Oy 121 tggcaaacgaggtctatatttctcactgtgtgtcacaagatattaaagaactgcag 180
Db 2271 tgaagagaagaagcatltaacttcaactgtgacaaatgacagcagcagaagaatgaagttg 2330
Oy 181 aacagatgagtcacatgacatatctta-----ctgaagaagaagacca 223
Db 2331 aattgaaatgcagctgtagtgaatccaaagtgaaatgtatgcttggaaatgcttccaa 2390
Oy 224 ctcaattcctctgcactctgtgaagagcatcaag-----tcag 261
Db 2391 atgattcaagatccagctcaataagaanaagacaactcgtagaggtgtcgtgagtcac 2450
Oy 262 acttattgacagagctgtaggaatccaccaatctaaagaagtaagttctctgagatct 321
Db 2451 aagcccaagacagaagcctagtagcacaagaagcctctgtagaagatataccccaagttc 2510
Oy 322 ctctatlaaaaatttlaaagttaaaccaagcctgaatggtccttltgtgtctgaggacat 381
Db 2511 ccttttgaggtattgaagcctaatttaacgaatgagcctatttgtgtgtgtgat 2570
Oy 382 tggcttctgtcttaaatggaactgttcatccagatatttccatcatcttggcaaaatca 441
Db 2571 ttgtgtccattlaaatgtaggagcctgcaaccagcatltgcaataatatttcaagaatca 2630
Oy 442 taacctgtgttgaataataagtaaaacacattaaag---catgagaagaattatt 498
Db 2631 taggggttttcaagaatgtgatgtcctggaacaaagaacagaatagtaactgtttt 2690
Oy 499 ccattgatactgtcatttgggtgtatttattgtgtgtgcattatcatgcaggatatt 558
Db 2691 cactatgttcttagcccttggaattatttcttattatcatttccctcagggtttca 2750

```

OY	559	tttaagcgagcgaggggaatttaacgaatagatlaagaaactgtgccttcaagcca	618
Db	2751	catttgcgcaaaactggaagatcctccacaacgcgtccgatacatgtgttcccgata	2810
OY	619	tgttatcatgagatattgcctgtgtttgatgaagaaggaaaacagcacaggatgttaca	678
Db	2811	tgctcagaagaagatgtagttgtgtttatgagacctaanaaacacacttgaagcttacta	2870
OY	679	caataatagcaatagataatagacaacaaattccaaggaagcaaacaggtttccaggatgtgcct	738
Db	2871	ccaggtctgcgaatgtagtgcctcaagttaaagggtgctatagttccagctgtcgttaa	2930
OY	739	taacacaaaatgcacaatacgaatgggaacttcagttatcattcctttatatatagatgg	798
Db	2931	ttaccagaataatagacaacacttctggacaggaataatatacttcatactatgtgttgc	2990
OY	799	agatgacattccctgattctcgtgafattgtctccagtaacttgcgtcgagacaggaatgatitaa	858
Db	2991	aactaacactgttaacctcttagcaattgttaaccatcattgaaatagcagaggtgtgttga	3050
OY	859	ccgcagcaatgactgtgatttgcacaacaagaatagaacgaagacttaagcagctgtcgga	918
Db	3051	tgaanaatgttctgttcggaacaagcctctgaagaagatagaagaagacatgaaggtgcgtggaga	3110
OY	919	tagcaactgaagcttcttgagaaatatacgttactatgttcatcaacaagggaaaagcct	978
Db	3111	tcgctactgaagaatagaaaacttcogaaacggtgtttctttgactcagagcagaagtt	3170
OY	979	tcgagcaaatgtctgaagagatgtcttcagactccaacacagaataactctcgaaagaacac	1038
Db	3171	ttgaacatagttatgtccagagatttgcaagttacatatacagaacactcttgaagaaagac	3230
OY	1039	agattattggaagctgtttagacatctgaacatgccttatataatttgcctatgcaacag	1098
Db	3231	acactcttggaaattacatttctctccaccaggaacagatgatatcttccctatgtgat	3290
OY	1099	ggttcgatttggagccatttaattcaagcttgaacgaatgaaacccagaggatgttca	1158
Db	3291	gttccgatttggagccttacttgttgcacataaactcatgaagctttagagatgttctgt	3350
OY	1159	tagttttactgcatttgcataaggagctatgtgcacatcggaanaagcctcgttttgcctc	1218
Db	3351	tagtatatttcaagctgtgttctcttcttgggccaatgcgtcggtggggaagttcaatttctc	3410
OY	1219	ctgaatatctcaaaagccaaatcgggggctgcgacatcgttttgccttcttggaanaagaac	1278
Db	3411	ctgacatactgcaaaagccaaaataatcagcagccacatcatgatgatitgaanaaaaccc	3470
OY	1279	caaatatagacagccgcagttcaagaagaaggaaaagccagacacatgtgaaggaaatttag	1338
Db	3471	ctttgattgacagctacgaacagcgacgttaatgctcggaacacatttgaagaatgtca	3530
OY	1339	agtttcgagaagctcctcttcttctatccatgctgcgccagatgttttcaactcctcgttgct	1398
Db	3531	catttggttgaagtgtatatacaactatccacccgcacccgaaacatccacatgcttcaaggac	3590
OY	1399	tatccctcagtatattgagcggaggaagaacagtagatttcttgaggagcgcgcgttggga	1458
Db	3591	tgagctcggaaggtgaagaaggccagacgcgtcggtctcgttggggagcagcgttgcgttggga	3650
OY	1459	aaagcattctgttcaactctcgcagagactttagaccccggttcaaggaacaaatgtctgt	1518
Db	3651	agagcaaaagtgtccagctcctctggaagcgttctctagacaccccttgcagggagaaggtgcgc	3710
OY	1519	ttgattggtgtgtagtcaaaaagaaattgaaatgtacagtggtctccgttccaaatagcaatcg	1578
Db	3711	ttgattgcaaaagaataaacccgactgaatgttctagtgctcccgagcaacacttgcgtacgtc	3770
OY	1579	ttctctaaagccctgtctcttcaacgtagcagcatttgcgtgaanaaacctcgcatgttgc	1638
Db	3771	tgctcccgagacatccctgttctttagctgcagcatgtgcgtgaagaacattgcccattgtgagca	3830

[illegible]

CC	RESULT	9
CC	AA294738	
PT	ID	AAZ94738.standard; cDNA; 4646 BP.
XX		
AC	AAZ94738;	
XX		
DT	01-AUG-2000	(first entry)
XX		
DE	Human ATP binding cassette ABCB1 (MDR1) cDNA.	
XX		
KW	ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;	
KW	atherosclerosis; lipid disorder; dyslipidemia; psoriasis;	
KM	lupus erythematosus; diagnosis; gene therapy; MDR1;	
KM	multidrug resistance; chromosome 7q21; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200018912-A2.	
XX		
PD	06-APR-2000.	
XX		
PF	21-SEP-1999; 99WO-EP06991.	
XX		
PR	25-SEP-1998; 98US-0101706.	
XX		
PA	(FARB ) BAYER AG.	
XX		
PI	Schmitz G, Klucken J;	
XX		
DR	WPI; 2000-293151/25.	
XX		
PT	Adenosine triphosphate binding proteins useful for identifying agents	
PT	for treating atherosclerosis and other inflammatory disorders -	
XX		
PS	Claim 9; Page 110-112; 154pp; English.	
XX		
CC	The present sequence is that of human ATP binding cassette	
CC	subfamily B protein ABCB1 cDNA. The cDNA was identified using a	
CC	differential display method in which monocytes from peripheral	
CC	blood were subjected to macrophage differentiation and cholesterol	
CC	loading with acetylated low density lipoproteins and subsequent	
CC	cloning with high density lipoprotein (HDL3) to identify	
CC	cholesterol sensitive genes. The gene maps to chromosome 7q21	
CC	and is also termed MDR1 (multidrug resistance). The invention	

CC provides cholesterol-sensitive ABC genes (see AA294734-63). These  
CC genes, and polypeptides encoded by them, can be used for diagnostic  
CC and therapeutic applications, and for biochemical or cell-based  
CC assays to screen for pharmacologically active modulator compounds  
CC useful for the treatment of lipid disorders, atherosclerosis or  
CC other inflammatory diseases such as psoriasis and lupus  
CC erythematosus.

XX Sequence 4646 BP: 1371 A; 892 C; 1129 G; 1254 T; 0 other:

Query Match 34.5% Score 712.4: DB 21; Length 4646;  
Best Local Similarity 60.1% Pred. No. 2e-181;  
Matches 1262; Conservative 0; Mismatches 796; Indels 42; Gaps 3;

QY 1 cgaagaaagtcgacatacaatcgtgtgtagcacacgacacttctactatcgaagtcag 60  
DB 2151 cgaagaaagtcgacacccattgtgatagctacatcgttctacagttcgaagtcg 2210  
QY 61 attgatgtgacccctaagatggaatgctgagcagagaaggaagacacatgctgaactaa 120  
DB 2211 acgtatcgtcgtgttcgacgtgacgtgacgtgacgtgacgtgacgtgacgtgacgt 2270  
QY 121 tggcaaaacgagtcataatattcactgtgtgacgtgacgtgacgtgacgtgacgtgacgt 180  
DB 2271 tgaagagaagaagcattacttcaactgtgtcacaatgacagacaggaatgagttg 2330  
QY 181 aacagatgagtcacatgacatattctta-----ctgaagaagaaccaa 223  
DB 2331 aattgaataatgcagctgatgataacagtgaaattgacgtgacgtgacgtgacgtgacgt 2390  
QY 224 ctcaactctctgcactctgtgaaagatcaga-----taag 261  
DB 2391 atgattcaagatccgacttaataaagaagaalcaactcgttagagatgctcgtgacac 2450  
QY 262 acttactgacaagcgtgagaatccaccacaaatcaagaagatgacgtcgtgacgtcgt 321  
DB 2451 aagcccaagacagaagcgtgacacaaagcgtcgtgacgtgacgtgacgtgacgtgacgt 2510  
QY 322 ctctataaaaaatttaagttaaaacagcctgaaagccttctgtgtcgtgagacat 381  
DB 2511 ccttttgagagatgacgttaaaatttaacgtgaatgacgtccttattgtgtgtgacat 2570  
QY 382 tggctcgttctaagtgaactgttccacgtatttccactccttcttgaataata 441  
DB 2571 ttgtgcataataaagtgaagcctgcaacacagcatttgcataatatttcaaaagata 2630  
QY 442 taacactgttggaaataatgataaacaacataaag---catgatgcagaatttatt 498  
DB 2631 taggggttttaacaagaatgtgacgtcgtgaacaaagaacagaagaatgtaactgtttt 2690  
QY 499 ccaatattcgtcatttctgggtgttatttgcctgtgcagttattcatgcaggaattat 558  
DB 2691 cactattgttctagcccttggaaattatttcttatttattcatttcttccacaggtttca 2750  
QY 559 tttaagagagcagaggaatatttaacgatgagtttaagaaactgtgcttcaagcca 618  
DB 2751 catttgcagaagcgtgagaagatcccaacaaacggtcgcgaataatgtgttccgaccca 2810  
QY 619 tgtatcacagatatctcgtgttctgtagtaaaagaaacagacagaggtcgtgacaa 678  
DB 2811 tgcctagacagatgt 2870  
QY 679 caatttgcataagataatgacaaatcaagagacaaaggttccagatgtgcgtct 738  
DB 2871 ccaggtcgcacatgctgctgcaagttaaaaggggtatagttccaggtcgtgttaa 2930  
QY 739 taacacaaatgcaactaacaatgaggaacttccgtatattccttattatattgtaggg 798  
DB 2931 taacacaaatgcaactaacaatcctgtgacaggaataattatattccttattatattgtaggg 2990  
QY 799 agatgacattcctgattctgagttatgtccagttacttgcgtgacaggaatgagtaaa 858

DB 2991 aactaacatgttactcttagcaattgtacccatcatttgcataatagcaggttctgaa 3050  
QY 859 ccgacgaatctactgatttccacaaagaataagcaagacttaagatgctgtgaaaga 918  
DB 3051 tgaataatgttctgtaaaacacactgaaagaataaagaagacttagaaggtcgtgaaaga 3110  
QY 919 tagcaactgaagccttggagaataatagctactatagttgtcatttaacaggaagaaagcct 978  
DB 3111 tgcctacccaagcaataaagaacttcgaacccgttcttctttagctcagaagcagaagt 3170  
QY 979 tcaagcaaatgtatgaagagatgcttcagactcaacacagaataatcctcgaagaagcac 1038  
DB 3171 ttgaacataatgtatgctcagaagtttgcaggtacacatacagaacacttctgagaagaagcac 3230  
QY 1039 agatatttggaaagcgttataatgacatgacgtccttataatttgcctatgagcag 1098  
DB 3231 acatcttggaaattacatttctcaccacccgagcaatgatbatttctcattgtcgtgat 3290  
QY 1099 gtttcgaatttggacccatttaattcaagctgagcgaatgaccccaagagcagctgttca 1158  
DB 3291 gtttcggttggagccacttgcgtgacacataaactcagacttgaagatgtctcgt 3350  
QY 1159 tagttttactgcaattgcatatagagctatggtccatcggaaaaagcgtcgttgcctc 1218  
DB 3351 tagtatcttgcagctgttcttctgttgcacatggtccgttgggaagtcagttcattgtctc 3410  
QY 1219 ctgaatattccaaagcacaatccggggtcgtgcacatcgttctgttctgttgaagaagac 1278  
DB 3411 ctgactatgcccacaaacataatcagccacacatacattatcatttgaaaaaaaccc 3470  
QY 1279 caaatatgacagccgcagctcagaagaaggaagaaagcagacatgcttgaaggaattag 1338  
DB 3471 cttgtattgacagccacagcagcaggaagccttaatgacgaacacatttgaagaagaaatgca 3530  
QY 1339 agtttcgaagaagtccttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1398  
DB 3531 catttggagagttgtatttcaactatccacccgcagacatcccaagtcgtcagggac 3590  
QY 1399 taccctcagatcttgcagagaagaagacagatgcatctgttggggagcagcgtgtgga 1458  
DB 3591 tgaagccttgaggttgaagaagggcagacgcttgccttctgttggcagcaggtgtgtgga 3650  
QY 1459 aagacacttctgtcaacttgcagagacattttagaccccggtgacagacagctgtgt 1518  
DB 3651 aagacacagtggtccagctcctggagcgttcttgcagaccccttgcagagggaagtgctgc 3710  
QY 1519 ttgacgtgtgtgaatgcaaaagaatgaaatgtaacgttgcgtccgtcccaatagcaatcg 1578  
DB 3711 ttgatgcaagaagaataaagcgacgtgaattgttcagttgctccgagacacacgttgcacgt 3770  
QY 1579 ttccccaagagccttgccttcaactcagacatgctcgtgagaacatcgctatgtgtgaca 1638  
DB 3771 tgtcccaagagccacatccctgttctgactgacgtgacgtgacgtgacgtgacgtgacgt 3830  
QY 1639 acagccgtgtgtgtcagatgagatgagatcaaaagacccaatgcagaatatccatt 1698  
DB 3831 acagccgtgtgtgtcagatgagatgagatcaaaagacccaatgcagaatatccatt 3890  
QY 1699 ctttatttgaaggttcccttgcagaaatacaacaacaagtttgactgaaagagcagc 1758  
DB 3891 ccttcatcagagtcactgccttaataatagcactaaagtaggagacaagaagacactcagc 3950  
QY 1759 ttcttgcggtcagagaagaacagactatgtcagaaggtccttctccaaaacccaata 1818  
DB 3951 tctctgtgtgcagagaacaacagcatttgcataagctgtgccttcttgaacagcctcata 4010  
QY 1819 ttcttgttgaatgagcacttgcagccctgcataatgaagcgtgagaaggtgtgtcagc 1878  
DB 4011 ttcttgccttggatgagacacgtcagccttgcataacagaagaaggtgtgtccaaag 4070  
QY 1879 atgccccttgaataaagcagagcaggaagagacatgctcagtagtgcacacagcctctctg 1938  
DB 4071 aagcccttgcagaagaagcagagagcagcagcctgtgattgtgtgtcctacccgctgtcca 4130





[illegible]

KM	retroviral: murine embryonic stem cell virus; MESV;	
KM	Moloney murine sarcoma virus; (MOMSV);	
KM	Friend murine leukemia virus; F-MuLV; ds.	
XX		
OS	Synthetic.	
XX		
PN	DEL1503952-A1.	
XX		
PD	14-MAR-1996.	
XX		
PF	07-FEB-1995; 95DE-1003952.	
XX		
PR	08-SEP-1994; 94DE-1431973.	
XX		
PA	(BOE) BOEHRINGER MANNHEIM GMBH.	
XX		
PI	Baum C, Osterlag W, Stocking-harbers C, Stockingharbers C;	
XX		
DR	WPI; 1996-152306/16.	
XX		
PT	Hybrid retroviral vectors - for gene transfer into haematopoietic	
PT	stem cells	
XX		
PS	Disclosure: Page 25-29; 42pp; German.	
XX		
CC	New hybrid vectors comprise (1) a leader region including the U5	
CC	region and RNA primer binding site of murine embryonic stem cell	
CC	virus (MESV) or Moloney murine sarcoma virus (MOMSV), and (2) a 3'-	
CC	LTR including the U3 and R regions of a Friend murine leukemia	
CC	virus (F-MuLV). The vectors are useful for ex-vivo or in-vivo gene	
CC	therapy. High levels of gene transfer can be achieved in	
CC	haematopoietic stem cells and their myeloid (non-lymphatic) progeny.	
CC	PSF1, PSF2, pSF3 and pM1 (sequences given in AAT13390-T13393) are	
CC	examples of such vectors.	
CC	Vector pSF-MDR (sequence given in AAT13394) is based on the	
CC	MESV vector R224.	
XX		
SQ	Sequence 6505 BP; 1719 A; 1506 C; 1606 G; 1674 T; 0 other;	
	Query Match 34.5%; Score 712.4; DB 17; Length 6505;	
	Best Local Similarity 60.1%; Pred. No. 2.3e-181;	
	Matches 1262; Conservative 0; Mismatches 796; Indels 42; Gaps	
QY	1 cggcgaaggtcgagctacatcgctgtagcacaccgactctctactatcgaaagtgcag 60	
DB	3543 ccgaagaaggtcgagaccacatgtagtgcacatcgcttgcctaagtcgtatgctg 3602	
QY	61 attgatgtgacccctaagaagatggaatgcgcggaagaagaagacacatgctgaactaa 120	
DB	3603 acgtcacgtcggttgcgttgatgtagtgcattggtggaagaagaatacatgtagactca 3662	
QY	121 tggcgaacagaggtcatatatttactctgtgatgtcacaagatatataaaagctgatg 180	
DB	3663 tgaagaagaagaagcatcttacttcaactctgtcacaatgtagacagcaagaagaattgaagtgc 3722	
QY	181 aacagctggagtcacatgacatatctca-----ctgaagaagaagcaca 223	
DB	3723 aattagaaaaatgtagctgtagaattcccaaaagtgaaatgtatgctctggaaatgcttccaa 3782	
QY	224 ctcaactctctgtagctctgtgaagaagcacaag-----ctgaag 261	
DB	3783 atgattcaagaatccagctcctaataagaanaagaatcaactcgttagaggtgtcgtgataaac 3842	
QY	262 aattcatgacaagcttgtaggaatcacccaactcaagaagaagaataagctctctgaagct 321	
DB	3843 aagcccaagaagcagaagcttagtaccacaagaagagcttgtagaagaatcaccctcagttt 3902	
QY	322 ctctattaaaaatttaagaagttaaaagaacgcggaatgggccttttggctggaggaat 381	
DB	3903 ccttttgaggagtagaagaagcctaatttaacggaatgggcctcatttggcttgggtgat 3962	
QY	382 tggctctcttccaatggaagctgttcacacagatatttccatcatccttgcaaaaatca 441	



Db 3963 ttgtgccaataaataagagcgccgcaaccagcattgcaataatttcaaatga 4022  
 Oy 442 taacacatttggaataatgataaaccacattaaag---cagatgcagaattatt 498  
 Db 4023 taggggtttcacagaagatctgactcctgaaacaaacagacagatgtaactgtttt 4082  
 Oy 499 ccatgatactgcatactttgtgtgttacttgccttgcagttatctcatgaggattat 558  
 Db 4083 cactattgttctgaccttggaattatttctttatcatatttcttccctcagggtttca 4142  
 Oy 559 ttacgagcagagcagggaataatttaacgaltgagatlaagacacttgcctcaagca 618  
 Db 4143 catttgcagaagctgagagctccaccagcgctccgatacatggtttccgatacca 4202  
 Oy 619 tgttatacagagatctccgtgttggatgtaaaagaaacagccagagagcttgacaa 678  
 Db 4203 tgcctacagagatctgtgtgttgaagccctaaacacacacttgagacattgacta 4262  
 Oy 679 caatatagccatagatataagacaacattcaaggagcagaggttccaggatgtgcct 738  
 Db 4263 ccagctcgcgaagatgctgtcctcaagltaaagggccataggtccaggtctgtctaa 4322  
 Oy 739 taacacaaatgcaactaacatgagccttcaagttatcatcttcttataatgagatgg 798  
 Db 4323 ttaccagaataagcaaatccttggacaggaataattatcatcttcatcatgtgtgc 4382  
 Oy 799 agatgacattcctgattcttgatgtgtcctcagatcttgcctgagcaagaatgtgaa 858  
 Db 4383 aactaacctgttactccttaagcaattgtaaccatcatagcaatgaggaattgttgaa 4442  
 Oy 859 ccgcagcaatgacttgatcttgcaacaaagaataagcaagaacttaagcactgtgaaaga 918  
 Db 4443 tgaataatgtctgacagcaagcactgaagaagataagaagaacttagaaggtctggaaaga 4502  
 Oy 919 tagcaactggaagcttggagaataacgtactatagtgctcattaaagaagaaaaagcct 978  
 Db 4503 tgcgtactggaagcaatagaanaactccgcaaccgtgttcttctgactaaggagcagaagt 4562  
 Oy 979 tcgagcaatgtatgaagagatgctcctcagactccaacagaaataccctcgaaagaagcac 1038  
 Db 4563 ttgacaatactatagctcagaggttgcaggtacccatcacagaacactcttgaggaagcac 4622  
 Oy 1039 agatatttgaagcctgtatgtatgtcatcagcactgcttataatttgcctacagcag 1098  
 Db 4623 acactccttggattacatttcccttaccacagcaatgatttcttccatgtcgtgat 4682  
 Oy 1099 ggttgcatttggagcctatttaattcaagctggaagatgaacccagagaggtgttca 1158  
 Db 4683 gttccggttggagcctacttgtgacacataacatcatgagcttltgagatgttctgt 4742  
 Oy 1159 tagtttactgcaattgcatatgagatgagtcacatgsgacatgsgaaaagcgtcttggctc 1218  
 Db 4743 tagtaatttcagcgtgtgtcttctgtgacagcggtggtgggaagtcagttcatgttgc 4802  
 Oy 1219 ctgaaatccaaagccaaatcgggggtgctgcatctgttgccttgttggaaaagaac 1278  
 Db 4803 ctgactatgccaagaacaaatcatcagcagccacatcatcatgacttcttgaataaaccc 4862  
 Oy 1279 caaataatgagacgcgcagatcaagaaggaagaagccagacacatgtgaaggaattag 1338  
 Db 4863 ctgtgattgagacgtacacgcaggaagccttaattgcgaacacacattggaagaatgtca 4922  
 Oy 1339 agttcgaagagctccttcttcatcactgtgcggcagaggtttcttccctcgtgtgc 1398  
 Db 4923 catttggagagtgatcttcaactatccaccgcagccgaatcccatccagtgctcaggac 4982  
 Oy 1399 tatccctcagatctgagcgaagaagacagatgattgttgggagcagcgctgtgga 1458  
 Db 4983 tgaagcttggaagtgagaagagcgagacgctgtctgtggtggaagcagtgctgtgga 5042  
 Oy 1459 aaagcaacttctgtcaacttctgcagagacttatgaccccggtcgaagacagtgctgt 1518

Db 5043 agagcagatgtgtccagctccttgagcggttctacgaccccttggcagggaagtgtctc 5102  
 Oy 1519 ttgagtgtgtgagtcacaagaatgtaagtgtacagtggtctcgttcccaatagaaatcg 1578  
 Db 5103 ttgagtcagaagaataaagcagctgtaattgtcgtgtcccgacacactgtggcactcg 5162  
 Oy 1579 ttcccaagagcctgtgtcttcaactgcaagatgtgcagaaactgcctcatgtgtgaca 1638  
 Db 5163 tgtcccaagagccactcctgttctgactcagcagatgtgcggaacacttgcctatggagaca 5222  
 Oy 1639 aacgcgtgtgtgccaatagatgagatcaagaagacccgaatgtgcacaaatccat 1698  
 Db 5223 aacgcgggtgtgtgtcagaggaagatcgtggaaggcagcaagaagacaaatatacatg 5282  
 Oy 1699 ctttattgaaggtctcctctgagaataacacacaaagtttgactgtgaagagcacag 1758  
 Db 5283 ccttcatgaatcactgtcctataataatagcactaaagttagagacaagaagaaactcagc 5342  
 Oy 1759 ttctcgtgcgcagaaanaagactagctattgtcgaaggcctctctccaaaacccaaa 1818  
 Db 5343 tctctgtgtgcagaaanaagcactgtccatagctcgtgtcccttgttagacagctcata 5402  
 Oy 1819 ttattgttgaatgagccacttcaagcctcagataatgacagltgagaaggtgtgtcagc 1878  
 Db 5403 ttgtccttgtatgaagcagcagctcagctgtgatacagaagaagtgaagaaggtgtccaag 5462  
 Oy 1879 atgcctctgataaagccagcggaagagatgcctgactgtgtcactcaacagctctctg 1938  
 Db 5463 aagcccttgaaanaagccagaaagagcgcagccttgatgtgtatgtcactccgctgtcca 5522  
 Oy 1939 caattcagaacgcagatttgaatgtgtctgcacaatgtgaaagataaaggaagcaagaa 1998  
 Db 5523 ccattccagaatgcagacttaatagtggtgttcagaaatgagcaggttcaaggagatgtgca 5582  
 Oy 1999 ctcatcaagagctcctgagaatcgcagacatatatttaagttgtgaaatgcagatcag 2058  
 Db 5583 cgcatacagagctgtgtgcacagaagaagcactatcttcaactgtgtcagtgccaagctg 5642

RESULT 12  
 AA249332 standard; cDNA: 3860 BP.  
 ID AA249332;  
 AC AA249332;  
 XX  
 DT 14-MAR-2000 (first entry)  
 XX  
 DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.  
 XX  
 KW Multidrug resistance: MDR-1; P-glycoprotein;  
 KW Transmembrane efflux pump; hematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassemia;  
 KW Gaucher's disease; sickle cell anemia; leukemia; ex vivo expansion;  
 KW cytokine; wild-type; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..3843  
 FT CDS  
 FT /\*tag= a  
 FT /product= "Human wild-type MDR-1 protein"  
 FT mutation  
 FT /\*tag= b  
 FT /note= "cDNA sequence of G185V human mutant MDR-1 given  
 in AA249332"  
 FT  
 FT  
 FT  
 XX  
 PN MO9961589-A2.  
 PD 02-DEC-1999.  
 XX  
 XX 27-MAY-1999; 99WO-US11825.  
 PF  
 XX

PR 28-MAY-1998; 98US-0086988.  
 XX (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX Sorrentino B, Bunting K;  
 XX MPI: 2000-072615/06.  
 DR P-PSDB: AAV58186.  
 XX Ex vivo expansion of hematopoietic stem cells transduced with a  
 PT sequence encoding human multidrug resistance-1, used for bone marrow  
 PR transplantation -  
 XX  
 PS Claim 10; Page 68-70; 113pp: English.  
 XX This sequence represents cDNA encoding human wild-type  
 CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane  
 CC efflux pump, responsible for the export of drugs from cells,  
 CC particularly cancer cells. Wild-type MDR-1 shows increased  
 CC resistance to etoposide and decreased resistance to vinca  
 CC alkaloids compared with a mutant form (AAV58187) where the Gly at  
 CC position 185 is replaced by Val. The invention relates to transducing  
 CC hematopoietic stem cells with nucleic acid encoding an MDR protein  
 CC and culturing the modified cells. The modified hematopoietic stem  
 CC cells are useful in bone marrow transplantation (to reconstitute  
 CC hematopoietic systems in patients who have undergone chemotherapy or  
 CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
 CC cells derived from hematopoietic stem cells, e.g., thalassemia,  
 CC Gaucher's disease, sickle cell anemia or leukemia. The modified  
 CC cells can also be used to identify factors involved in regulating  
 CC proliferation and differentiation in hematopoietic stem cells.  
 CC Hematopoietic stem cells that express MDR-1 will be protected against  
 CC chemotherapeutic agents, so can be engrafted while the patient is  
 CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells  
 CC provides sufficient cells to permit standard biochemical analysis.  
 CC Overexpression of MDR-1 allows cytokine-driven expansion of  
 CC hematopoietic stem cells by at least 10-fold compared with a maximum  
 CC of 4-fold in known procedures.  
 XX  
 SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;  
 Query Match 34.4%; Score 710.8; DB 21; Length 3860;  
 Best Local Similarity 60.0%; Pred. No. 5e-181;  
 Matches 1261; Conservative 0; Mismatches 797; Indels 42; Gaps 3;

QY 382 tggcttcgttctaataatggaactgttcacacagatatttcacatcatttgcaaaaat 441  
 Db 2147 ttgttgccatttaaatggaaggcctgcgaacagcatttgcaataatttcgaagaat 2206  
 QY 442 taaccatgttgaataatgataaaccacattaaag--cattatgagaattatt 498  
 Db 2207 taggggttttacaagaatttgatgatacctgcgaacaaacagacagatagtaactt 2266  
 QY 499 ccatgatalctgcattttggtgtatttgccttgcattatcattgcagggat 558  
 Db 2267 caattgtttcctagcccttggaattatttcttattacattttcctcaagttca 2326  
 QY 559 tttaggcgacgaggggaattttaacgatgataagacattgaccttcaagcca 618  
 Db 2327 catttgcgaacgttggaagatccctcacaacgctgcgatacgttttcgcatca 2386  
 QY 619 ttattacagatattgctgtgttgatgaaaggaacagcagcaggttcacaa 678  
 Db 2387 tgcctcagacagatgtaggttggtttgatgacccctaaacacacattgagcttga 2446  
 QY 679 caattatgacatagatatagcacaaatccaaggagcaacaggttcagagtgccgt 738  
 Db 2447 ccaggtcgcgaatgtagtgcctcaagttaaagggtatagttccagcttgcgttaa 2506  
 QY 739 taacacaaatgcacataacatgaggaacttcagttatcattccttatatagatgg 798  
 Db 2507 ttaccagaatatagaacaaacttggagacaggaataattatcattcattatagtt 2566  
 QY 799 agatgacattccgattccgagatgctgcacagtaacttccggaaggaatgtagaa 858  
 Db 2567 aactaacactgttactccttagcaattgtatccctcatttgcataagagagtgtagaa 2626  
 QY 859 ccgacgaatgactggaattggccacaagaataagcaaaccttaagcatgctggaaga 918  
 Db 2627 tgaanaatgttctggaacaagcactgaaagataagaagaactagagtgctcgggaaga 2686  
 QY 919 tagcaactgaagctttggagaataatagctactatagttcatataacaaaggaaagc 978  
 Db 2687 tgcctactgaagaatagaaataacttcggaacggttcttcttgactcaggaacgaagt 2746  
 QY 979 tcgagcaaatgtatgagaagatgcttcagacatacacaagaataactcctggaagaagc 1038  
 Db 2747 tgaacatatgtatgctcagagtttgcaggtacacatacagaataacttggaggaagc 2806  
 QY 1039 agatattggaagctgtatgcatltagcatcagcatgcttataatatttgctatgcagcag 1098  
 Db 2807 aactccttggaattaatatttctcctcaccaggaacatgatatatttctcctatgcgtat 2866  
 QY 1099 ggttcgatttggagcctatttaattcaagctgagcagaatgacccagagggcatgttca 1158  
 Db 2867 gttccggttggagcctactcttggccatataactcagaaacttggagagtgcttctgt 2926  
 QY 1159 tagttttactggaattgcaatgtagagctatggcctatggaanaagcctcgtttggcct 1218  
 Db 2927 tagtatttccaagctgttgccttgcgtgcacatggcgttgggacagtcagttcaatgtc 2986  
 QY 1219 ctgaatatcccaagcacaatcggggctgagcagatgttggcttggttggaagaagaac 1278  
 Db 2987 ctgacatagcgaacgaacaaatatacagcagccacatcatcatgatcatgtgaaataacc 3046  
 QY 1279 caaatatgacagcgcagtcgaagaaggaanaagccagacacatgtgaaaggaattag 1338  
 Db 3047 ctgttgattgacagctacacagcaggaagcctaatatgcgaacacatgtgaaaggaattca 3106  
 QY 1339 agtttcgagaagctccttctcctatcactatgtagccagatgtttcaatcctcgtgctc 1398  
 Db 3107 catttggtagagttatcatcaactatccacccagacggacacacccatgcttcaggagc 3166  
 QY 1399 tatccctcagtatggagcagaagaagcagtaacatttgggtggaagcagcgttgggga 1458  
 Db 3167 tgagccttgtaggtggaagagggccagacgctgctcgttgggtggaagcagtggtggtgga 3226

```

OY 1459 aagacactctgtcaactctctgcagagacttattgaccccgtagcaagaatgctcgt 1518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3227 agagcaacgctgtgtccacaggaagatctgtgagcggtctcaagacccttgccagggaaatgctgcg 3286
OY 1519 ttgatgctgtgagcaaaagaatgaatgacagtgagctccgtctcccaatagaatcag 1578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3287 ttgatgcaagaataaagaagcactgaatgtctcagctgagcagacacgttgagcatcg 3346
OY 1579 ttccccaagaagcgtgtgtcttcaactcagcagatgctgtagaagactgcgtatgtagca 1638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3347 tgcaccaagagcccaactcctgttgcagcagcattgctggaagaacatgacctgtagagaca 3406
OY 1639 aacagccgctgtgagccatagatgagatcaaaagaagccgcaaatccacaatalccatt 1698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3407 aacagccggtgtgtgtccacaggaagatcgtgagagcagcaagaagagccacaataactc 3466
OY 1699 ctttatgaaagcttccctgtagaaatacaacacacaagcttgagactgaagagcacagc 1758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3467 ccttcacagctcactgcttaataataatagcactaaagtagagagacaagaactcagc 3526
OY 1759 ttctcggcgccagaaacaagaactagctattgcaaggcgctcttccaaaaaaccccaaa 1818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3527 tctcggctgagccagaaacaacacgcatctgcacatagcctcgccctgttagacagcctcata 3586
OY 1819 tttaattgtgagtagagccacttcagccctgcagataatgacagtgagagagtggttcagc 1878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3587 ttctgcttctgtgagtagaagcagctcgtcgtgatacagaaagtgaaaaagtggttccagc 3646
OY 1879 atgccttgataaagccagagcagagacatgacctagtgctcaactcaagagctcctcg 1938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3647 aagcccttgacaagaagcagagaaagcgacccgacatgctgattgctcacccgctgtcca 3706
OY 1939 caattccagaacgacatttgatgagtggtctgcagacatgagaaagtaagaagacaaggaa 1998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3707 ccattccagaatgacacttaataatgagtggttccagaaatgagcagagtcagagcatggca 3766
OY 1999 ctcatcaagagcctcctgagaaatcagacacataatttaagttagatgaatgcacagctag 2058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3767 cgcatcagcagctgctgtgcacagaaagcactatttttcaatgtgtcagtgctccagctg 3826

```

```

RESULT 13
AAZ49333
ID AAZ49333 standard; CDNA: 3860 BP.
XX
AC AAZ49333;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human G185V mutant multidrug resistance-1 (MDR-1) CDNA.
XX
KW Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW cytokine; mutant; ds.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key 1.3843
FT CDS /tag= a
FT /product= "Human G185V mutant MDR-1 protein"
XX
XX WO9961589-A2.
XX
XX 02-DEC-1999.
XX
XX 27-MAY-1999; 99WO-US11825.
XX
XX 28-MAY-1998; 98US-0086988.
PR

```

```

XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Sorrentino B, Bunting K;
XX
DR WPI: 2000-072615/06.
DR P-PSDB: AAY58187.
XX
XX Ex vivo expansion of hematopoietic stem cells transduced with a
PT sequence encoding human multidrug resistance-1, used for bone marrow
PT transplantation -
XX
XX Example 1: Page 79-82; 113pp; English.
XX
XX This sequence represents CDNA encoding human G185V mutant multidrug
XX resistance protein MDR-1, where the Gly residue at position 185
XX of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
XX transmembrane efflux pump, responsible for the export of drugs from
XX cells, particularly cancer cells. The wild-type MDR-1 shows increased
XX resistance to etoposide and decreased resistance to vinca alkaloids
XX compared with the G185V mutant. The invention relates to transducing
XX haematopoietic stem cells with nucleic acid encoding an MDR protein
XX and culturing the modified cells. The modified haematopoietic stem
XX cells are useful in bone marrow transplantation (to reconstitute
XX haematopoietic systems in patients who have undergone chemotherapy or
XX radiation therapy) and in ex vivo gene therapy of genetic defects in
XX cells derived from haematopoietic stem cells, e.g., thalassaemia,
XX Gaucher's disease, sickle cell anaemia or leukaemia. The modified
XX cells can also be used to identify factors involved in regulating
XX proliferation and differentiation in haematopoietic stem cells.
XX Haematopoietic stem cells that express MDR-1 will be protected against
XX chemotherapeutic agents, so can be engrafted while the patient is
XX undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
XX provides sufficient cells to permit standard biochemical analysis.
XX Overexpression of MDR-1 allows cytokine-driven expansion of
XX haematopoietic stem cells by at least 10-fold compared with a maximum
XX of 4-fold in known procedures.
XX
XX Sequence 3860 BP: 1135 A; 746 C; 956 G; 1023 T; 0 other;
XX

```

```

Query Match 34.4%; Score 710.8; DB 21; Length 3860;
Best Local Similarity 60.0%; Pred. No. 5e-181;
Matches 1261; Conservative 0; Mismatches 797; Indels 42; Gaps 3;
OY 1 cgagcaaaagctcgcagctacatcgtggtgtagacaccgactttctactattggaagtgcag 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1727 ccagaaaaggtgcgaccacacatltgatagcctcagcttgcacagttcgtatgctg 1786
OY 61 attgatgtgacccctaaagatatggaatcgtgcgagaaagagacatgctgaactaa 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1787 acgtcatcgcgtggttgcgtatgtagtgcacatgtgagaaagaaatcatgatgaactca 1846
OY 121 tggcaaaagaggtctatattatcacttgatgtcagacagatatataaaagctgag 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1847 tgaagagaaagagcattacttcaactgttcacacatgcagacagcaggaataagagtg 1906
OY 181 aacagatgagtcacatgcatacttca-----ctgaaagaaagaccac 223
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1907 aattgaaatgcagctgatgaatccaaagtgaattgatgccttggaaatgttcttca 1966
OY 224 ctcaactctcgtcacctcgtggaagagcatcaag-----tcag 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1967 atgattcaagatcagctcttaataaagaaagatacaactcgtagagagtgctcgtagatac 2026
OY 262 acttcattgacaaggtctgaggaatccaccacacttaagagataagcttctctgaagctc 321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2027 aagcccaagacgaagagcttgtagtaccagaagagcgtctgtagaagtagtactccagctt 2086
OY 322 ctctattaaaaatttaagttaacaagcctgaagccttctgtgttctgtgggaact 381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2087 cctttagagatatagcacttaactgaatgaatgaatgaatgaatgaatgaatgaatgaatga 2146

```

Db	3227	agagacagcgtgttcacagctcttcyggcggttctacagccctcttgcagggaaagtctcgc	3286
Oy	1519	ttgatggtgtgatgcacaagaatgaattacagtggtccgttcccaatagatcacg	1578
Db	3287	ttgatggtcaagaataataaagcgactcgtaattgltcagtggtctccgagacaacactcgggcatcg	3346
Oy	1579	ttcctcaagaagcgtgtgctctctcaacttcgaagcatctgtcgaacacatcgctctatgtagca	1638
Db	3347	tgctccagagagcccatcctgtttagctcgcagatctgtcgaagacatltgcattgtagagca	3406
Oy	1639	acagccgtgtgtgttcattagatgcataaagaagccgcaaatgcagcaaatatccatt	1698
Db	3407	acagccgtgtgtgttcacagagagatcgttgaaggcagcaaaagagcccaacatcacg	3466
Oy	1699	ctttattgaaaggtcttccttgagaataacaacacagttctgactggaaggagcacagc	1758
Db	3467	cccttaccgaagtcactgccttaataataatagcaaccaagtaagagacaaagaaactcagc	3526
Oy	1759	ttctcgcgcgcagacaacaaagactagatctatgcgaaggctctctccaaaaccnaaa	1818
Db	3527	tctctggtggccagagaacaacgcatctgcacatgcctcgtgcttgttagacagctcata	3586
Oy	1819	tttatgttggatgtaggcacactcagccctcgataatgcagtgagaaggtgtgtcagc	1878
Db	3587	ttttgcttttgatgaagcagcaagctcagctctgtatcacagaaggtgaaaggtgtccaa	3646
Oy	1879	atgcctctgataaagccagagaaaggaagaaactgtgctgtgtgactctacacagcctctg	1938
Db	3647	aagcccttgacaaagccagagaaagggccgacactgcatctgtgtctcacagcctgtcca	3706
Oy	1939	caattccagaacgcagatttgatagtggtcttcacacatggaagaataaagagacaagaa	1998
Db	3707	ccatctcagaatgcagacttaataagtgtgttcagaagaagcagatccaagagagcttgca	3766
Oy	1999	ctcatcaagagctcctcctgagaatcgcagacataatttaagttagtgaaatgcacagcag	2058
Db	3767	cgcatcagcagtgctgtgcacagaaagcatctatttccaatgtgtcagctcagctg	3826
RESULT 14			
AAZ88973	AAZ88973 standard; DNA; 3988 BP.		
AC	AAZ88973;		
DT	30-MAY-2000 (first entry)		
XX	Human MDR-1 DNA.		
XX	Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;		
KW	stomach; colon; ulcerative colitis; pouchitis; large intestine; human;		
KW	inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy		
XX	ileum; anti-ulcer; anti-inflammatory; cytoslatic; MDR-1; ss.		
OS	Homo sapiens.		
XX	WO200007577-A2.		
PN	17-FEB-2000.		
XX	06-AUG-1999; 99WO-DE02426.		
XX	06-AUG-1998; 98DE-1035526.		
PR	15-DEC-1998; 98DE-1057750.		
XX	(STRE//) STREMMEL W.		
PA	Stremmel W.		
XX	WPI: 2000-195439/17.		
DR	Using phosphatidylcholine having mucous membrane protective activity to		



QY	1639	acagccggtgtgtgtgcacattagatgagatcaaaagaagcccgcaatgtcagTcaaatatcatt	1658
Db	3473	acagccggtgtgtgtgtgcacaggaagagattgttgaaggcgacgaagaaggtgccaaactacatg	3532
QY	1669	cttttatctgaaggtctccccgtgagaaataaacaacaagctgtgactggaagaagacacgc	1758
Db	3533	cccttcaacgaggtacacgtgccttaataataatagaacctaaaggtgaagaaacaaggaactccgc	3552
QY	1759	tcttcctgcgagccaggaacaagaactgactgtgcgaaggtctcttcccaaaaccca	1818
Db	3593	tctctgtgtgcagaaacaacagatgtgcatagtccgtgcctctgtttagaagctcata	3652
QY	1819	tcttatctgtctgtagtgagccactcgaagccctgcgaataatgacagtgtagaaggttgctcaagc	1878
Db	3653	tttctgctctttgtgtagtagaagccacgtctgcagctctgtgatacagaagaagtgtgaaaggttctcaag	3712
QY	1879	atgccccttgtagtaagccagagacggtggaagatcagctgctgtgtctacatacaggtctctgt	1938
Db	3713	aagcccctgcagaaagccagagaaaggtccgacacccgtcaatctgtgtctgtccacgcctgtgtcca	3772
QY	1939	caatctcagacagcaagaattgtatagtgtgtctctgcacatctgtgaagaagaataaaggaacagagaa	1998
Db	3773	ccatccagaaatgacagacttaataatgtgtgtcttcagaaatgtgcagagtcgaagagactgtgca	3832
QY	1999	ctcatcaagagctccctctgagaataatcgagacatatattttaagttagtgaatgcacagttcac	2058
Db	3833	cgatctcagcagctgtctgtgcacagaaagatcatatttccaattgtctcaggtgtgtccagagctgt	3892

RESULT	15
AAF86127	
ID	AAF86127 standard; cDNA; 4186 BP

DT	25-JUN-2001	(first entry)
XX		
DE	Cynomologous monkey P-glycoprotein CDNA.	
XX		
KM	Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;	
KW	efflux pump; ss.	
OS	Macaca fascicularis.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	100..3942
FT		/tag= a
FT		/product= "PGP"
XX		/note= "P-glycoprotein"
XX		
PN	WO200123565-A1.	
XX		
PD	05-APR-2001.	
XX		
PE	28-SEP-2000; 2000WO-US26592.	
XX		
PR	28-SEP-1999; 99US-0156921.	
PR	12-OCT-1999; 99US-0158818.	
XX		
PA	(GENT-) GENTEST CORP.	
XX		
PI	Stocker PJ, Steinel-Crespi DR, Crespi CL;	
XX		
DR	WPI: 2001-316136/33.	
XX		
DR	P-PSDB; AAB81064.	
PT		
PT	Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein	
PT	(PGP) and homologous PGP polypeptides are useful for predicting	
PT	biavailability of compound and increasing PGP transporter activity in	
PT	cell	
XX		
XX		
PS	Example 1; Page 51-57; 84pp; English.	

XX This invention relates to a polynucleotide sequence encoding a  
CC cynomolgous monkey P-glycoprotein (Pgp), and an allelic variant of the  
CC Pgp protein. Pgp, also known as multidrug transporter, MDR1 is a member  
CC of the ABC transporter superfamily. The enzyme serves as an efflux pump  
CC exporting small molecules across the cell membrane. The invention  
CC includes a cynomolgous monkey (*Macaca fascicularis*) Pgp coding sequence  
CC and protein, and also that of an allelic variant. The Pgp polynucleotide  
CC sequence is useful for increasing Pgp transporter activity in a cell.  
CC Antisense sequences of the cDNA are useful for inhibiting Pgp transport  
CC activity in a mammalian cell. They may also be used for increasing the  
CC bioavailability of a drug. The present sequence represents cDNA encoding  
CC cynomolgous monkey P-glycoprotein.  
XX  
XX Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;  
SQ

Query Match	34.4%;	Score 710.8;	DB 22;	Length 4186;
Best Local Similarity	60.0%;	Pred. No. 5.1e-181;		
Matches 1261;	Conservative	0;	Mismatches 797;	Indels 42; Gaps 3

OY	1	cgagcaaaagtcgacatacaatcgtgtgtgacacccgacttctctactatcctgaagtgaag	60
Db	1826	ccagaaaaagtcgcgacccacccattgtatagatgcatactcgttctccacggttcgaattgcg	1885
OY	61	attgattgtgacccctaagaatgacgtatcgtgcgagaaagaaagacaatgctgtgacaa	120
Db	1886	acgtcaltccgctgtcttcgagatgagtgatcattgttgagaaagaaatcattgatgagccca	1945
OY	121	tggcnaaacgagtcctatattatcactgt-----gatgcacagga	163
Db	1946	tgaagagaaagcatttacttcaactcttgcaaatgcaaatgcagacaggaattgaaattgt	2005
OY	164	tattaaaaagctgatacaagatgtgacatgacatacttacttgaaagaagaccaa	223
Db	2006	aattagaanaatgcagctgtatgnaatccaanaatgaaattgatacctcttgnaatgtcttca	2065
OY	224	ctcaactccctgcgacctgtgaaagagcatcaag-----tcag	261
Db	2066	atgatctcagatccagctctataagaanaagatccactctgtagaggtgtcgtgatatca	2125
OY	262	actcattgaaagaagctgaaagaaatcccaacctaaagaataagtcctcctgaagct	321
Db	2126	aagccaaagacagaagaagcttagtaccaaagaagctctgtgataagaaatbaactccagatt	2185
OY	332	ctctattaaaaatttaagaattacaacagctgaaatgaccttctgtgtctcgtggagacat	381
Db	2186	ccttttgagagattatgaagctaaatttaactagtagtgacctatttctgtgtgtgtat	2245
OY	382	tggtctcgttcttaaggaagctgtcatccagattttccacbaactcttgaaaaata	441
Db	2246	cttgccattataataatgagagctctgnaacacagcatttgcagtaaatatttcaaaagatta	2305
OY	442	taaacatgttctggaataatgataaacaacacatgaaag---catgtagcagaataattat	498
Db	2306	taaggatttttaacagaanaatgatagtatgcccgaacaacaaagacagaatagtaactgtttt	2365
OY	499	ccatgataatcgtcaatttgggtgttaattgtctcttgacgtattatcatgacagggattat	558
Db	2366	cactattgtttctagctcccttggaattgtttcttttattacatttcttccctcagggcttca	2425
OY	559	tttaacgacagagcgaggaaattttaacgtatgatatgaacactctgacctcagaagctca	618
Db	2426	catttgcgaagaagcgcggagagatccccaacagcgctccgaatacatalggtttccgaacca	2485
OY	619	tgtataatcagatatgcctgtgtttgataaagaagaaacagcacaagagatgcgaaca	678
Db	2486	tgtccagacagagatgtgagctgttttgatgacccctaanaacaacacactgtgagatatgacta	2545
OY	679	caatatattagccatagatatagacaaaattcaagagacaagctccacagattgagctct	738
Db	2546	ccagagctgcacagatgagctgtcccaagttttaaaggagctatagattcccaagctgtcataa	2605







THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 15:11:32 ; Search time 14771.4 Seconds  
(without alignments)  
2307.374 Million cell updates/sec

Title: US-09-873-409-9

Perfect score: 2066  
Sequence: 1 cgcagcaagatgcgactaca.....atgcacagtcagtcagtcga 2066

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_ov:\*  
5: gb\_om:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rod:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	739.8	35.8	3924	6 AX024455	AX024455 Sequence
2	739.8	35.8	3924	9 HUMMDR3	M23234 Human membr
3	718.2	34.8	4018	5 GGA9799	AJ009799 Gallus ga
4	712.4	34.5	2726	6 I33621	I33621 Sequence 1
5	712.4	34.5	4646	6 I49610	I49610 Sequence 2
6	712.4	34.5	4646	6 HUMMDR1	M14758 Homo sapien
7	712.4	34.5	4669	6 I08557	I08557 Sequence 3
8	712.4	34.5	6505	6 AR028671	AR028671 Sequence
9	712.4	34.5	9318	6 AR028672	AR028672 Sequence
10	710.8	34.4	3988	6 AX024454	AX024454 Sequence
11	710.8	34.4	4186	6 AX108654	AX108654 Sequence
12	710.8	34.4	4192	6 AF016535	AF016535 Homo sapi
13	710.8	34.4	4195	6 AX108656	AX108656 Sequence
14	710.8	34.4	4264	6 AR051647	AR051647 Sequence
15	710.8	34.4	4264	6 AR051650	AR051650 Sequence
16	710.8	34.4	4669	6 AR055785	AR055785 Sequence
17	710.8	34.4	8630	6 AX012321	AX012321 Sequence
18	709.2	34.3	4378	6 E02326	E02326 Multidrug r
19	709.2	34.3	8630	6 AX012320	AX012320 Sequence
20	704.4	34.1	4669	6 AR091275	AR091275 Sequence
21	702.4	34.0	4279	6 AX105057	AX105057 Sequence
22	702.4	34.0	4279	6 AX105078	AX105078 Sequence
23	702.4	34.0	4279	6 AX105080	AX105080 Sequence
24	700.8	33.9	4045	12 AF269224	AF269224 Synthetic
25	700.8	33.9	4279	6 AX105082	AX105082 Sequence
26	700.8	33.9	4317	4 AF045016	AF045016 Canis fam
27	700.8	33.9	4317	6 AX105059	AX105059 Sequence
28	698	33.8	3987	10 CRUPGP1185	M59254 Chinese ham
29	698	33.8	4296	10 CRUPGP1	M60040 C.griusens p
30	698	33.8	4304	10 CRUPGP1165	M59253 Chinese ham
31	693	33.5	4280	10 CRUPGP11	M60041 C.griusens p
32	690.2	33.4	4084	10 MUSMDRA	J03398 Mouse mdr g
33	685.4	33.2	3905	10 CRUPGP111	M60042 C.griusens p
34	685.2	33.2	3858	10 OAU78609	U78609 Ovis aries
35	682.4	33.0	2409	10 CRUPGP2A	M17896 Chinese ham
36	672	32.5	2080	10 CRUPGP1A	M17897 Chinese ham
37	659.6	31.9	4927	10 AF257746	AF257746 Rattus no
38	657	31.8	4323	10 AF286167	AF286167 Rattus no
39	651.6	31.5	3912	10 RATPGLYCO	L15079 Rat p-glyco
40	646.2	31.3	4298	10 MUSMDR	M14757 Mouse multi
41	645	31.2	4233	6 AR123273	AR123273 Sequence
42	637.2	30.8	3682	10 MOSPELYIA	M24417 Mouse phosp
43	637.2	30.8	4924	10 MUSMDRA	M33581 Mouse p-gly
44	635.6	30.8	4356	10 MUSMDRAA	M30697 Mouse multi
45	627.2	30.4	4390	5 XLU17608	U17608 Xenopus lae

## ALIGNMENTS

RESULT 1  
AX024455 LOCUS AX024455 3924 bp DNA PAT 15-SEP-2000  
DEFINITION Sequence 2 from Patent DE19857750.  
ACCESSION AX024455  
VERSION AX024455.1 GI:10184623

KEYWORDS human.  
SOURCE ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 3924)  
AUTHORS Stremmel,W.  
JOURNAL Patent: DE 19857750-A 2 24-FEB-2000;

FEATURES  
SOURCE STREMMEL WOLFGANG (DE)  
Location/Qualifiers  
1..3924  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="MDR3/menschliche MDR-Genfamilie"



Db 3805 AGCTGCTGCACAGAAAGCATCTATTTTCAATGCTCAGTGTCCAGCTGGG 3857

RESULT 2

LOCUS HUMMDR3 3924 bp mRNA PRI 11-JUN-1993

DEFINITION Human membrane glycoprotein P (mdr3) mRNA, complete cds.

ACCESSION M23234

VERSION M23234.1 GI:187501

KEYWORDS P-glycoprotein; membrane glycoprotein.

SOURCE Human liver, cDNA to mRNA, clone 3.27.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 3924)

AUTHORS van der Bliek, A.M., Koolman, P.M., Schneider, C. and Borst, P.A.

TITLE Sequence of mdr3 cDNA encoding a human P-glycoprotein

JOURNAL Gene 71, 401-411 (1988)

MEDLINE 89138016

COMMENT Draft entry and computer-readable sequence for [1] kindly provided by P. Borst, 21-MAR-1989.

FEATURES

source

1..3924

/organism="Homo sapiens"

/db\_xref="taxon:9606"

<1..3924

/product="mdr3 mRNA"

33..3872

/note="P-glycoprotein"

/codon\_start=1

/protein\_id="AAA36207.1"

/db\_xref="GI:307181"

/translation="MDLEAAKNGTAMRPTSAEGDFELICSSKOKKKTKTYKMICVLT  
LEFRSDWQDKLFMSLGTITMAIAGHSGLEPLMIIVFEMDKFVDTPGNSFPVNFSLT  
LNPKILDEEMTRAYVYSGAGVLAVALVAVTIOGSPVLESTIAINICGRNVMDEIKKAVKE  
EIGMEDINDTELETRDDISIKSEGIDKVMFOAVAFPAFIFGFIKMLTL  
VIMASPLTGLSAAMWAKILSAFSDKELAAKAGAAVEALGARTVYIAFGOKEL  
ERYOKHLENAEIGIKKASINISMGITFLIYALAFVGTSLVISEKTTINAM  
TVFSTILIGAFSVGOAPCIDAFANAKAAVTFIDINNPIDISEKHPSIDIK  
NLEFNDVHSPSRANVKILKGLNKVOSGOVALVAGSSCGKSTVOLIORLYPDE  
GTINIGDQDIRNFNNYLRLEIGVYVSOEPLVSTIAINICGRNVMDEIKKAVKE  
ANAYEFIMKLPOKFDPLVGERGALSGGOKORIAIARALVRNPKILLDEATSDTE  
SEAEVOALDKAREGRTIVIAHRLSTIRNDVIAFGDVGVEGSSKMSQCKRGVY  
FKLVNMQTSQIOSEFELNDEKATMAAPNGKSRIFRSTONLNKNSQCKQSLD  
VETDLEANPPVSVFLKVLNKTETWPFVVGTCVAINGGLOPFSVSETIAIFG  
PGDAVAKQCKNIFSLIFLGLIISFEFFLOGTFPKAGELTIRLSMAKAMLR  
DMSWDDHKNSTRLANDAOVAGATGRLALIONINLGTGIIISFIYGMOL  
TLILAAYPIIAVSGIVEMKILAGNAKDKKLEAGKATAPAINIRIVVSLTOERK  
FESMVEKLIGPYRNSVOKAHYIGITFISQAFMFSVAGCRRGPAIVLNGHMERD  
VILVPSAIFGAVALGHSFAPDPAKAKLSAHLFMLERPPLIDSYSESLKDKK  
EGNTTFNEVFNPTIRANVPVLOGLSLEKKQOTLALVSSCGKSTVOLIERFYD  
LAGTVLIDGOEAKKLNVOMLRQGLIVSOEPLIFDCISAENTAYVDSNIVVSDIYS  
AAKANIHPLEITLPHKYEYTRGDKTOLSGGOKORIAIARALIRDPDILDEATSA  
LDTESEKIVOEALDKAREGRTIVIAHRLSTIRNDVIAFGDVGVEGSSKMSQCKRGVY  
FKLVNMQTSQIOSEFELNDEKATMAAPNGKSRIFRSTONLNKNSQCKQSLD  
KGIYSMSVSOAGTONTL"

BASE COUNT 1145 a 790 c 977 g 1012 t

ORIGIN

Query Match 35.8%; Score 739.8; DB 9; Length 3924;  
Best Local Similarity 60.6%; Pred. No. 8.4e-147;  
Matches 1268; Conservative 0; Mismatches 792; Indels 33; Gaps 2;

QY 1 cgagcaaaagtcgactacatcgtagtagcacaccgacttctactatctgaagtcag 60  
1 11 1111 1111 11 1111111111111111 1111 1111 111111  
Db 1765 CCAGAGAAAGCGGACCATGTGATAGACACCGCATGCTACGGTCCGAATGCA 1824  
1 11 1111 1111 11 1111111111111111 1111 1111 111111  
QY 61 atttgatgtgacccctaaagatggaatgctgtagcgagaaaggaacatgctgaacta 120  
1 11 1111 1111 11 11111111 11 1111 1111 11 111111  
Db 1825 ATGTCTATTCCTGGGTTTGAGAGTAGAGTAATGTGGAGCAAGGACCAAGCACTGA 1884  
1 11 1111 1111 11 111111 11 1111 1111 11 111111  
QY 121 tggcaaaacgagtgctatctatctcactgtgagt----- 154  
1 11 1111 1111 11 111111 11 1111 1111 11 111111

Db 1885 TGAAGAAGAGGGGTGTACTTCTTAACCTTGTCACATGCAGACATGCAGAAAGCCAGATCC 1944

QY 155 -gtcacagatatataaaagctgatatgaacagatgtagtcaatgacatatctcagcaaa 213  
1 1111 11 1111 1111 11 111111111111 11 1111 1111 111111111111

Db 1945 AGTCAGAAAGAAATTGAACTAAATGATGAAAAGGCTCCACATTAAGTGGCCCCAAATGGCT 2004

QY 214 gaaagacacactactcctctgcaactcgtgtagagagcatcaagtcagacttgcagaa 273  
1 11 1111 1111 11 11111111 11 1111 1111 11 111111111111

Db 2005 GGAATCTCGCTATTATTTAGGCATTCTACTAGAAAACCTTTAAATTCACAAATGTGTC 2064

QY 274 aggc-----tgaagatccacccaactcaaaagataagctctcgaagtcctcat 327  
1 11 1111 1111 11 1111 1111 11 11111111 11 1111 1111 111111

Db 2065 AGAAGAGCCTTGATGTGGAACCGATGACCTTACAGCAAAATGCGCACAGCTCCTTTC 2124

QY 328 taaaaatttaagttaaacaaagcctgaatgagccttctggtctgtaggagcatgagctt 387  
1 11 1111 1111 11 11111111 11 11111111 11 1111 1111 111111

Db 2125 TGAAGCTCTGAAACGAAATTAAMACAGAAATGGCCCTACTTGTGCGGGGAAAGTATGTG 2184

QY 388 ctgcttaagatgaactctgcatcagatatttccatcatcctctgcaaaaataataacca 447  
1 11 1111 1111 11 111111 11 111111 1111 11 1111 1111 111111

Db 2185 CCATTGCCAATGGGGGCTTACCGCGCATTTTCAGTCAATATCTCAGAGATATAGCGA 2244

QY 448 tgttggaaataaataaaacacataaaagcatgacagaatattatccatgatat 507  
1 111111 1111 11 1111 1111 1111 1111 11 1111 1111 111111

Db 2245 TTTTGTGACACAGCGCATGTGATGACGTAGACAGCAGAAAGTGCACATATTCCTTTGATT 2304

QY 508 tgcattcttgagtgctattctgcttgcagtgatattcagaggaatttcaagcca 567  
1 11 1111 1111 1111 11 1111 1111 11 1111 1111 11 111111

Db 2305 TCTTATTTCTGGGAATTAATTTCTTTTACTTCTTCTCCACGAGGTTTCACGTTTGGGA 2364

QY 568 gaggcgggaaatttaagatgagatlaaagacacttgagcctcaagccatcttatc 627  
1 1111 1111 1111 11 1111 1111 11 1111 1111 1111 1111 111111

Db 2365 AACCTGCGGATCTTACACAGAAAGCTGCGTCAATGGCTTTTAAAGCAATCTTAAGAC 2424

QY 628 agaatatccttgcttctgtagtgaagaaagaaacagcaggaagcttgcacaacataatag 687  
1 1111 1111 1111 1111 11 111111 11 1111 1111 11 111111

Db 2425 AGGACATGAGCTGTTTGAATGACCATTAATAAACAGTGTGTCACCTTTCACAAAGCTTG 2484

QY 688 ccaataatagcacacaaatcaagagcaacaggttccagattgtagcttcaacacaa 747  
1 1111 1111 1111 1111 1111 1111 1111 11 1111 1111 111111

Db 2485 CCACATATGCTGCCCAAGTCCAGAGACCCACAGAACACAGGTTGGCTTTAATGGACAGA 2544

QY 748 atgcaactaactaggacttcaagttatatacttcccttataatgtagggagaatagat 807  
1 1111 1111 1111 1111 1111 1111 1111 11 1111 1111 111111

Db 2545 ATATACCTTAACCTTGAACCTGATTAATCAATATTAATCAATGCTGACAGTTAACCC 2604

QY 808 tccctgattcgaatattgctccagtaacttgcgtgcaaggaatgatatcaaccgcagcaa 867  
1 11 1111 1111 1111 1111 1111 1111 1111 11 1111 1111 111111

Db 2605 TATTGCTATTAGCAGATTGTTCCAAATTAATTCGTGTGCAGAAATGTTGAATGAATGCT 2664

QY 868 tgaactgatttgcacaacaaagataaagaacacttaagcatgctggaagaatagcaactg 927  
1 1111 1111 1111 1111 1111 1111 1111 11 11111111 1111 111111

Db 2665 TGCGTGAAGATGCCAAAGAGATTAATAAGACATGCAAGCTGCTGGAAATGTCACACAG 2724

QY 928 aagcttggagaaatatacgtactatagctcttaatacaaggaagaaagcccttgagcaaa 987  
1 11 1111 1111 1111 1111 1111 1111 1111 11 1111 1111 111111

Db 2725 AGGCAATAGAAAATATTTAGACAGTTGTGCTTTGACCCACGAAAGAAAATTTGAATCA 2784

QY 988 tgaataagagatgcttccagactcaacacagaaataactcctgaagaaagcagatatatg 1047  
1 1111 1111 11 1111 1111 1111 1111 1111 1111 1111 1111 111111

Db 2785 TGTATGTGAAAATTTGATAGACCTTACAGAAATTCGTGCAAGGACACATCATGATG 2844

QY 1048 gaaactgttatgcatcgaacgaatgacttatataatttgcctatgcaagaggttgcag 1107  
1 11 1111 11 1111 1111 1111 1111 1111 1111 1111 111111

Db 2845 GAATTTACTTTTACTATCTCAAGACATTTATGATTTTTCCTTATGCGGTTGTGTTTCGAT 2904

QY 1108 ttgagcatatttaatttaacagctgagcgaatgagcccaagggcagtgatcaatagatttta 1167  
1 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 111111

Db 2905 TTGTGTGATATCTCATTTGGAATGACATATGCGCTTACAGAGATGTTATTCGTGTGTTT 2964

QY 1168 ctgcaattgcatactagagctatgagcacaacggaagaaacgcgcttctgagcctcgaataat 1227  
1 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 111111

Db 2965 CTGCAATTTGTAATTTGGTGCAGTGTGCTGTAGACATGCGCAATTCATTTGCTCGAGACTATG 3024

OY	1228	ccaaagccaatctggggctctgcgcatctcgttgcctctgttggaaaaaagaacccaatatag	1287
Db	3025	CTTAAAGCTAAACCTGCTGACGCCCACTTATTTCATGCTGTTTGAAGAACAACACTCTGTGATTG	3084
OY	1288	acagccgcagctcaagaagaaggaaaaagccagacacatgctgaaggaaatttagagttcggag	1347
Db	3085	ACAGCTACAGTGAAGAGGGGGCTGAAGCCTGTATTAATTTTGAAGGAATAATATAACATTTTAATG	3144
OY	1348	aagtcctcttcttctatccatctgtccagatglttccatctccgttgcgtatccctca	1407
Db	3145	AAGTGTGTTTAACTATCTCCACCCGAGCAAACTGTCAGTCCCTTCAGGGGGCTGAGCCTGG	3204
OY	1408	gtattgacgcggaaagaacagtagcatcttctggggagcaagcgcctctgggaaagaacatc	1467
Db	3205	AGGTAAAGAAAAGCCACAGCACTAGACCTCGTGGGGCAGCAGTGGCTGTGGGAGAGCAACGG	3264
OY	1468	ctgttcaactctctgaaagagacttatagaaccccgctgcagaagacaagtctgtttatggtg	1527
Db	3265	TGTTCCACTCTCTGGAGGGTTTACGACACCCCTTGGGGGAGACAGTGCTTTCGATGATGTC	3324
OY	1528	tgtatgcacaaagaattgagtatcaagtggtctccgttcccaatatgcatctgttccctaag	1587
Db	3325	AAGAAGCAAAAGAAATCATATGTCCAGTGGCTCAAGCTCAAGCTCGGATCGTGTCTCAGG	3384
OY	1588	agccctgtcctctcaacgcgcagcatctgtctgaacatctgcctatctgtctgacacagccgtg	1647
Db	3385	AGCCTATCTTATTGACCTGCGACGACTGTGCCGAGAAFTATTTGCTATTGAGACACAGCCGGG	3444
OY	1648	tgtgtgcattatgatagatcaaaagaagccgcaaatctgcagcaaatatccatctttatctg	1707
Db	3445	TTTGTATCCAGAGATGCAAAATTTGTAGTGCAGCCAAAGCTCCACAACTATCATCTTTCATCG	3504
OY	1708	aagctctccctcgaggaataacacaacaagaagcttgaaactgaaagggcagagttcttcggcg	1767
Db	3505	AGACCTTACCCCAACAAATATGAAACAAGAGTGGAGATAAAGGAGACTTCAGCTTCACAGAG	3564
OY	1768	gcacgaacaaagaactagctatctgcaagggctctcttccaaaaagcccaaaatttatctt	1827
Db	3565	GTCAAAACAGAGGATTGCTATTGGCCGAGGCCCTCATCAGCAACCTCAATTCCTCTGT	3624
OY	1828	tgtatgagtgccaactcagccctcgatataatgacagtgtgaaggtgtgttcagacatgccttg	1887
Db	3625	TGGATGAAGCTACATCACTACTCTGTGATCTGAAAGTGAAGAAAGTTGTCCAAACACCTTGG	3684
OY	1888	ataaagccacgaagcggaaagacatgcctcagtggtgcactcacaagggctctctgcaattaga	1947
Db	3685	ACAAAAGCCAGAAAGGCCGCACTGCAATTGTGATTGCTCACCGCTGTCCACCATCCAGA	3744
OY	1948	acgcagattctgagtgtgtcttcgacaaatggaagataaagaaagaaagaaatcatcaag	2007
Db	3745	ATGCGAGCTTATATAGTGTGTTTTCAGATGCGAAGTCAAGAGACATATGCGACGATCAGC	3804
OY	2008	agctctctagaacatccgaaacatatatttaagtttagtgaatgacacagtcatg	2060
Db	3805	AGCTCTCTGGCACAGAAAGCATCTATTTTTCAAATGGTCAAGTCTCAGGCTGGG	3857

RESULT	3			
GGA9799				
LOCUS	GGA9799	4018 bp	mRNA	VRT 03-MAR-1999
DEFINITION	Gallus gallus mRNA for ABC transporter protein.			
ACCESSION	AJ009799			
VERSION	AJ009799.1	GI:3355756		
KEYWORDS	ABC transporter protein; cmt1 gene; P-glycoprotein. chicken.			
SOURCE	ORGANISM Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus; 1 (bases 1 to 4018) Edelmann,H.M.L. Direct Submission			
REFERENCE				
AUTHORS				
TITLE				

JOURNAL	Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular Genetics, Biocenter and University of Vienna, Dr. Bohr-Gasse 9/2 A-1030 Vienna, A-1030 Vienna, AUSTRIA
REFERENCE	2 (bases 1 to 4018)
AUTHORS	Edelmann, H.M., Duchek, P., Rosenthal, F.E., Foger, N., Glackin, C., Kane, S.E. and Kuchler, K.
TITLE	Cmrl1, a chicken P-glycoprotein, confers multidrug resistance and interacts with estradiol
JOURNAL MEDLINE	Biol. Chem. 380 (2), 231-241 (1999)
FEATURES	99209605
source	Location/Qualifiers
gene	1..4018
CDS	/organism="Gallus gallus" /db_xref="taxon:9031" /tissue_type="intestine" 109..3975 /gene="cmrl1" 109..3975 /note="ABC transporter protein"
	/codon_start=1 /product="P-glycoprotein" /protein_id="CAA08835.1" /db_xref="gi:3355757" /db_xref="SPRMBL:093437" /translation="MHSEDEKQKITYVGNITYEINSNODDPEDEKGGKKKKKKPOMVSPALFRISCTDKLMTFESLAIAGHSTPLAAMIIFGDMTSEYTSGMTNTIGNSGLNMSADVENFKEEMETRYAYSYSAIAVAALVAVYIQTSEMTLAAGRVKRIKEFEFHAIKROEIGWEDKVDNAGELNTRLRIDVSEIGDKIGFIIQSETFNFGIFGFIIRKMLTVILAVSPVIGLSAALMAKILIAFDPEQAAVAKAGVAEEVLSAVRTVIAAGGKEIKIRYHNKLNEDARKIRIKRAITSNIGSAGFLIIVASVLAIPYGTLLILANFYSIGVILVPSYVLIGASIGCTAPSIAPANAKRAAIAIRNIIDNEPIDISYSPAGKRPDHKGNLEFONVEFNFPYSRDVLILGLNLKVCGQVALVVGSGGSKSTVOLIORFYDEREGTITIDGODKLSLVNRYLRETIIGNVODPVLEPATTIAINITYGREDYMEIERAKENANAYDFIMKRLPKFETTVGEGAGMSGGOKRIAIARLVNHPYIILDEATYSAULTSESYVOALDKARBCRTYVVAHRLSTVRNADLAVFESGVITBOGNSOLIEKRTIKLVNMOTIENDEPSEKSENAVSYKRGSGSNIDESIKRELRRGSPRRSMKKKCPGNPDNDEKSSPDDELPPVSEFLAKMLKKNWPYFVAGTEFATYNGALOPAFSVYFKSELIGFSETDQVLRKSNLYSLLELAGIISFFFEVQGFNFGAGETLIMKLEEFMAFMAERODMAWDPDKNSTGALTRILANDASQVAKGATGVRLLIOMNLNLGTGIIISLVYGMQLTLILAAPPYIIVAGMIEMKMLGAKDKILEANGLIAGEALEINIRTYVASLITRKRPELIVGEHLVPRNSVKRKAHIFGCFSLQOMPEFTYFAGCEPQAYLVVNGHIEKTYFLVPSAVYFQGMALGQSSSPADYAKAKISAAHIFVLPNRPYIDSVRENGDKPEKFGSGNTRIKDYKRYPNRPPEKYLIGQNLAVVEKETLALVGGSGCKSTIVYQLERFYPDLSEGLVFDIDDKTLTIOMLRSHIVISEQPLEFTFYIAENIYAGDNRSEVSEHETISAKKASIHISFDISLEKNTLRHGDGTQLSGOKORIAIARLIRKPDQIILDEATYSAULTSEKITYOELADKARGCRTYVIAHRLSTIQNADKIAVIONGVIBQDGTGHOQLAEKGFYSILNVGSGSCNM"
mat_peptide	112..3972
	/gene="cmrl1" /note="P-glycoprotein" /product="ABC transporter protein"
BASE COUNT	1288 a 758 c 911 g 1061 t
ORIGIN	
Query Match	34.8%; Score 718.2; DB 5; Length 4018;
Best Local Similarity	60.6%; Pred. No. 3.2e-142;
Matches 1270; Conservative	0; Mismatches 783; Indels 42; Gaps 4;
OY	1 cgagcaaaagtcggaactcaatcgtggtgtagcagcacacgcgaacttctactatctcgaagtgcag 60
Db	1859 CAAAGGAGGAGGAGCCACGCGTTGTAGTAGGCTCAACGCTCTATCAACACGATCCGAAATGAG 1918
OY	61 attgattgtgaacctaaagatgtaagtcgctggcggaagaaaggaacacatctgtgaactaa 120
Db	1919 ATCTTATCGCTGTGTTTGAAGTGGAGTTATTCACGAACAAGGAATCATTTCTCAACTGA 1978
OY	121 tggcaaacgagtcatactatatactcaatctgtgattgcacagagataataaaagctgat- 179
Db	1979 TTGAATAAAGGAGATCTACTATAACTCTGTTAAACATGCAGACATTAAGAAATCGAAGATC 2038
OY	180 -----gaacagatggagtcacatgacatatttctactgaaagaagaagccaac 224

Db	2039	CATCATCAGAAAAAAGTGA	AAATGCAGTGTGTCAAAAAGATGCATCTCAATCAATC	2098
QY	225	tcactccctctgcac	-----tgtgaagcatcaagtcaga-cttcattgcaag	276
Db	2099	TGGATGAATCCCTAAAAA	AAAAAAGAAFTAGAAAGAGATCACTACAGAGAAAGCATGAAAAAGC	2158
QY	277	ctgaagaaatccacccaat	ctaaagataagt-----cttcctaa	318
Db	2159	CAGGAGAACCTATATGAT	ACTGATGAGAAAGTGTACCTGATGAAGAATTACTCCAG	2218
QY	319	tctctctataaanaaltt	aaaglttaaacaagccctgaatgacctttgtgtctggga	378
Db	2219	TTTTCATTTCTGAAACT	TATGAACATTTAAACAAAAACGATATGGCCATACTTGTAGCTGGAA	2278
QY	379	catgagctctgtctc	aaatgagacggttcacacgatttccatcttcgaatcttgcaaaa	438
Db	2279	CTTTCTGTGCAATTTGT	CAATTTGAGCTTTTACACCTTGCAATTTGAGTATATTCACAGAA	2338
QY	439	ttataaccaatgttgg	aaataatgataaacaacatlaaagcaatgcatgcaagaaattat	498
Db	2339	TTATATAGGATTTT	TTTTCAGAACTGATCAGAAAGTTTGAAGAAAAAGCACTTATATTT	2398
QY	499	ccatgatatctgcata	tttttggtgttatttgcctttgtcagttatttcaatgcaaggattat	558
Db	2399	CACCTGCTGTTTTT	TAGCACCCTGGATCATTTTCTTTTACTTTCTTTGTTCCAGGGTTTTG	2458
QY	559	tttaagcgagacag	ggggaatttttaacagatgatatlaaacactgtgtcccttaagca	618
Db	2459	CATTTTGGCAAACTG	AGAAATTTCTTACATGAAAGCTTGCAATTCATGCACTTTTAAACAA	2518
QY	619	tgttatalcaagata	tttgcctggttttgatgataaaggaaacagcaacagagaggttgcacaa	678
Db	2519	TGCTGAGACAGGAC	ATGAGCTGCTGTGTGATGATCCCTAAAAATATGACATCGAGCACTGACAA	2578
QY	679	caatatgaacata	ggatataagcaaatccaagagcaaacagttccagatttggtct	738
Db	2579	CANAGACTTCTAT	TATGATGCTTCACAAAGTAAAGAGCACTGTTGTTAGATTGGCATTTAA	2638
QY	739	taacacaaatgca	actaatacatgagactttcagttatcatcttccattatataatagatggg	798
Db	2639	TTGGCCCAAAACAT	AGATTAACCTGGGAACGTGAATTTTATATCTTTAGTTTATNGGTGGC	2698
QY	799	agatgacatctc	tgattcttgagattgtccagttactgtgcgtgcaaggaatgatgaaa	858
Db	2699	ACCTGACCTGTTAC	TTACTTACCTGTTGTACCAATTCATGTGAGTGGCGAGGAATGATTGAA	2758
QY	859	ccgcagcaatgac	tgtgattgttccaacaaagaaatagaagaacttaagcatgctgcygaaga	918
Db	2759	TGAAAGATGTTGG	CGTGGCATGCAAAAAAAGATATAAATATAGAGCTGCGAAGCTGCGAGAAAGA	2818
QY	919	tatgcaactgaa	gcttttggagaataatacgttactatagttgltcataacaaggaaaaagcct	978
Db	2819	TTGGCCACACAAAC	CAATAGAAAAATTTTCGAACCTGTCTACTATTGACCTCGGAAAAAAGAT	2878
QY	979	tcgagcaaatgta	tgaagagatgcttcagactcagactcaacaagaagaataactcgaagaagaac	1038
Db	2879	TTGAGCTCATGTA	CGAGAACATTTTCTGTGTACCAATACAAATTTCTGTAAAGAGGCGC	2938
QY	1039	agattattgaa	gctgttlatgcatltaagcatgcatgctttatataatttgcctatgcaacag	1098
Db	2939	ACATATATTTGG	TTTTTTTCCCTTTCACAAGCAATGATGTTTTTTTACTATAGCTGCGCT	2998
QY	1099	ggtttcgaattg	gagcctatttaattccaagctgagcaagaatgaccccaagaggcatgttca	1158
Db	2999	GTATTCCGCTTT	GGAGCTATCTTGTGTGTTAATNGAACATATAGATACAAAAACGTGTTTTT	3058
QY	1159	tatgttttaact	gcaattgatalatgagcatagtcacatcgaaaaaacgctgtttgtctc	1218
Db	3059	TAGTGTGTTCA	CTGTTGATTTTGGTGCTATAGCATTAGACAAACACAGCTCTTTTGAC	3118
QY	1219	ctgaatatctca	aagccaatcgagggtctgcgcatctgttgcctttggtgaaagaac	1278

Db	3119	CAGACTATGCCAAGGCCAAGATATTCAGACGACCCACTTGTTTGTGCTTTTAAATAGATAC	31178
QY	1279	caatatagacagccgcgaatcagaagaaggaataaacgcacagacacatgtaagggaatttag	1338
Db	3179	CTCCAAATAGACAGTTTACCCGGAGAGTGGAGAGAAACCTTAATAATTTGGGGGAAACACA	3238
QY	1339	agtttcgagaagctctctctctcattccatcgaatgctgcgcgaatgttttcacccctcgctgc	1398
Db	3239	GGATTCMAAGATGTAAATTATTTACTATTCCAACCGACACAGAGTCAAAATCTCCCAAGGCT	3298
QY	1399	tatccctcagatattgagcgaagaagaacagtagcaattgttggggagcagcgctgttggga	1458
Db	3299	TGAATCTAGCAGTGTGAAATAAAGGAGAAACGCTTGCCCTTGTGTGTAGCAGATGTCTGAA	3358
QY	1459	aaagcactctcgttcaactctctgcagagacttatatccccgttgcgaaggaacaagtcgct	1518
Db	3359	AGAGCACTGTTGTTCACTTACTTGAGNGATTTTATATCTCACTCAGTGGAGAAATTTGCT	3418
QY	1519	ttagatggtgtgagatgcacaagaatltgaatgtacagtggtcttcgttcccaatatgaatcgc	1578
Db	3419	TTGATGACATGTAGTACAAAGACACTTAATATCCAGTGGCTGAGATCTCAGATCGGTATC	3478
QY	1579	ttccctcaagagcctgtgtcttcttcaactgtcagatctgtcgtgaagaatcgctctatgttgca	1638
Db	3479	TCTACACAGACCCCATCTCTTTCGACTTCCACATCTCCTTAACAAACATCGGTACGGGGACA	3538
QY	1639	aacagcctgtgtgtccatattagatgatcagaagaagccgcgaatgcgaataatcatcat	1698
Db	3539	ACAGTCCGGAGGTTCCCATGAGAGAAATCTTTAGTCCACGAGAAAGCAGCAGTATTCTATT	3598
QY	1699	ctttatttgaagtgctctccctcgtgagaatatcacacacaagaatltgactgaagaagagcacagc	1758
Db	3599	CCTTCACTTGTACTCTCTCCGCAAAAAATACAAATCTGTTAGAGACAAAGGAAACACAGC	3658
QY	1759	ttttgtgcgcagagaacaagaactagctatgtgcgaagcgtcttctccaaaaccccaaa	1818
Db	3659	TTTCTGGGGGCAAAACAAACGATTTGTCTATTTGCTTGAGCCTCTTATAGCTAAGCCCCAGA	3718
QY	1819	ttttattgtttgaagagccacttcacagcctctcgatataatcacagttgaagaagtggttcagc	1878
Db	3719	TTTGTACTACTGATGAAGACTACATCTGCTTGGACACGAAAGCGAAAGATTGTCCAGG	3778
QY	1879	atgcctctgtataaagccagagcagcggaagagaatagcttagtgttaactcaacagcgtctctg	1938
Db	3779	AAGCACTGGATAAAGCCAGAAAGGTGCGACCTCGCATGTGATGCTCACCGCTGTCTCA	3838
QY	1939	caattcgaagcagcagattcatatagtggttcgtcacaaatggaagaataaagaagacaagaa	1998
Db	3839	CCATCCAGAAATGCTGCACAAAGATTGTGTGTATCCAGAAAGCGCAAGGTCAATAGACGAAGGA	3898
QY	1999	cttcatcaagagctctctgtagaagaatcagacaatatatttaagttagttgaatgcaca	2053
Db	3899	CTCACACGACGCTCTCTGCTGAAAAAAGGCTTCTACTACTCTGCTGTCATATGTTCA	3953
RESULT 4			
LOCUS	133621	2726 bp	DNA
DEFINITION	Sequence 1 from patent US 5593840.		
ACCESSION	133621		PAT
VERSION	133621.1	GI:1824412	
KEYWORDS	133621.1 GI:1824412		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2726)		
TITLE	Bhatnagar,S.K., George,A.L. Jr. and Nazarenko, I.		
JOURNAL	Amplification of nucleic acid sequences		
FEATURES	Patent: US 5593840-A 1 14-JAN-1997;		
SOURCE	Location/Qualifiers		
BASE COUNT	818 A	538 C	642 G 728 T
/organism="unknown"			



RESULT 5  
149610 LOCUS 149610 4646 bp DNA PAT 07-OCT-1997  
DEFINITION Sequence 2 from patent US 5641508.  
ACCESSION 149610  
VERSION 149610.1 GI:2471830  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4646)  
AUTHORS Li, L. and Lishko, V. K.  
TITLE Method for delivering melanin to hair follicles  
JOURNAL Patent: US 5641508-A 2 24-JUN-1997;  
FEATURES  
source 1. 4646  
BASE COUNT 1371 a 892 c 1129 g 1254 t  
ORIGIN  
Query Match 34.5%; Score 712.4; DB 6; Length 4646;  
Best Local Similarity 60.1%; Pred. No. 5.4e-141;  
Matches 1262; Conservative 0; Mismatches 796; Indels 42; Gaps 3:  
QY 1 cgagcaagctcgagactcaacatcgtagagcacaccgacttctaactatcgaaatgcag 60  
Db 2151 CCAGAAAAGGTCGGACCCACCATTTGATAGCTCATCGTTTCTACAGTCCGTAATGCTG 2210  
QY 61 attgatgtgacacctaaagatggaatgcatgctgagcgagaagaagacacatgctgaactaa 120  
Db 2211 ACGTATGCGCTGGTTTCATGTGAGTGAATGTCATGTGAGAAAGAAATCATGTATGAACTCA 2270  
QY 121 tggcaaaacgagctcatatcatatcatctgtgatgtcacagatatlaaaagctgagatg 180  
Db 2271 TGAAGAGAAAGGCGATTACTTCAACACTGTCTACACATGCAGACACAGAAATGAAAGTTG 2330  
QY 181 aacagatggagatgcaatgacatactca-----ctgaagaagaagacaa 223  
Db 2331 AATTGAAAATATGCACCTATGATGATCCAAAGTGAATTTGATCCCTTGGAAATGCTTCA 2390  
QY 224 ctacactctctgcacctctgtgaagaagatcaag-----ctag 261  
Db 2391 ATGATTCAGAGATCCAGCTCTAATGAAGAAAAGATCAACTCGTAGAGTCCCTGATCAG 2450  
QY 262 actcatgacagaagctgagaaatccaccgaatctaaagagaataagctctctgaagctc 321  
Db 2451 AAGCCCAAGACAGAAAGCTTACTACCAAGAGAGCTCTGGATGAAGTATACCTCCAGTTT 2510  
QY 322 ctctaataaaatttaagttaaaacagcctgaaatggcctctgtggtcttcggggaacat 381  
Db 2511 CCTTTGGAGAGTATATGAAGTAAATTAACGTAAATGGCTTAATTTTGTGGTGTAT 2570  
QY 382 tggctctgctctaaatggaactgttcacagatatttccatcatcttgcacaaatla 441  
Db 2571 TTTGGCCATTATATATATGAGAGCTGCAACACGACATTTGGCAATATATTTTCAAGATTA 2630  
QY 442 taaccatgtttggaataatgataaaaccacatlaag----catgctgacagaatttct 498  
Db 2631 TAGGGGTTTTTACAAAGATTTGATGATCCTGAAACAAACAGACAGATAGTAATCTGTTTT 2690  
QY 499 ccatgatatctgcatcttgggtgttatctgtctgtcagtatcatctgacagggatcat 558  
Db 2691 CACTATGTTTCTACGCCCTTGGATTAATTTCTTTATATACATTTTCCCTGGGGTTTCA 2750  
QY 559 tttaagcgacagcgagggaaatttaacgatagataaagaactgtgaccttcaagca 618  
Db 2751 CATTTGGCAAAAGCTGGAAGATCTTACCAACAGGCTCCGATACATGATGTTTCCATCCA 2810  
QY 619 tgttatcagaatatgctgtgtgtgataaagaaacagcaagagagctgacaa 678  
||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2811 TGCTCAGACAGATGTGAGTTGTTGATGACCCCTAAACACACCATGAGCATTTGACTA 2870  
QY 679 caatttagccataatatagacaaatccaagagcaaacaggttccaggaatggcgtct 738  
Db 2871 CCAGGCTGCGCAATATGCTGTCTCAAGTTAAAGGGGCTATAGTTCCAGGCTGTGTAA 2930  
QY 739 taacacaaatgcaactaacatgagacattcagttatcatcttcatatataatgagatg 798  
Db 2931 TTACCCAGATATATACCAATCTTTGGACAGCAATATATATATCTTTCATCTTATGCTGGC 2990  
QY 799 agatgacatctcctgattctgagatagtcctccagtaactgctcgctgacagaatgatgaa 858  
Db 2991 AACTAACACTGTACTCTTACGCAATTTGTAACCATCATTCATTCGATACGAGAGTTGTTGAA 3050  
QY 859 ccgcagcaatgactgatttgcacaaagaagataagcaagaacttaagcatgctgaaaga 918  
Db 3051 TGAATATGTTGCTGTGACACAGCATGAAAGATAAAGAAAGACTAAGAGTGTGGGAAGA 3110  
QY 919 tagcaactgagccttggagaatataagctgactatagtgctcatlaacaagggaaaagcct 978  
Db 3111 TCGTACTGAACCAATAGAAAACCTTCCGAACCGTTGTTCTTTGACTCAGAGCATGAAGT 3170  
QY 979 tcgagcaaatgatatgaagatgcttcagagctcaacacagaataatcctcgaaagaagcac 1038  
Db 3171 TTGAACATATATGCTGTCTCAGACTTTGGAGTACCATACGAAGAACTTTGAGGAAGCAC 3230  
QY 1039 agattatggaaagctgttatgcatctcagcaatgctcttatatatatttgcctatgcaag 1098  
Db 3231 ACATCTTGGAAATATACATTTTCTTCCACCAAGCAATGATATATTTTTCATATGCTGAT 3290  
QY 1099 ggttcgatttggagcctatttlaactaaagctgagcgaaatggccccagaagcagatgca 1158  
Db 3291 GTTTCGGTTTGGAGCTACTTGTGTGGCACATTAATCTCATGTGAGCTTTGAGATGTTCTGT 3350  
QY 1159 tagtttacttcaattgcatatagagctatgagctacatcgaaacacgctcgtttgtgctc 1218  
Db 3351 TAGTATTTTACACTGTGTGTGTTGGTGGCCATGGCCGTGGGCAAGTACATTTGTTGCTC 3410  
QY 1219 ctgaatattccaagccaaatcgggggcttgcgcatcgtttgcttctgttggaaaagaac 1278  
Db 3411 CTGACTATGCCCCAAACCAAAATATCAGCGACGACATCATGTGATCATTTGAAAACCC 3470  
QY 1279 caatatagacagccgagctgaagaaggaagaaacccagacacatgtagaggaaattag 1338  
Db 3471 CTTTATTTGACAGCTACACGACGGAAGGCTTAATGCCGAACCATTTGGAAGGAATGTGA 3530  
QY 1339 agttcgaagaagctccttctctatccatgctgcgcccaaatgtttcaaccctcgctgct 1398  
Db 3531 CATTTGGTGAAGTGTGATTTCAACTATCCACCCGACGACATCCACAGCTTTCAGAGAC 3590  
QY 1399 tatccctcagtaattgagcgaagaagacagtagcatltgttgggagcagcggctgttggaa 1458  
Db 3591 TGAGCCTGAGAGTGAAGAAAGAGGCGACAGACGCTGCTGTGGAGACAGATGCTGGGGA 3650  
QY 1459 aagaacatctgttcaactctgcagagacatttaagaaccgcttgcagaagaagctgct 1518  
Db 3651 AGAGCAGACTGTGCTCAGCTCTGGAGCGGTTTACAGACCCCTTTGGCAGGGAAGTGTTC 3710  
QY 1519 ttgatgtgtgagatgcaaaagaattgaatgtacagtgctgctgtccccaatlagcaatg 1578  
Db 3711 TTGATGGCAAGAAATTAAGGCACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATGC 3770  
QY 1579 ttccctcaagagcctgtgctcttcaactcagcagatctgtgagaacatcgctctatgttga 1638  
Db 3771 TGTCACAGAGAGCCATWCTGTTTGTACTGACAGATTGCTGAGAACTATTCCTATGAGACA 3830  
QY 1639 acaagcgggtgtgtgcataatagatgagatcaagaagcgcgcaaatcgacgaataatccatt 1698  
Db 3831 ACAGCCGGGTGTGTACAGAGAGATGCTGAGGGAAGAGGACCAAGGCTCAATCAAG 3890  
QY 1699 ctttatttgaagtgctccctgagaataatacaacaagaattgtaactgaagaagagcagc 1758  
Db 3891 CCTTATGAGTACACTGCTCTATTAATATATAGCACTRAAGTATAGGAGACAAAGAACTCAGC 3950  
||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||





QY 224 ctcaactccctcgtactctgtgaagagcatcaag-----tcag 261  
Db 2391 ATGATTTACAGATTCAGTCTTAATAAGAAAAAGATCACTCGTAGAGTGTCCGTGATCAC 2450  
QY 262 acctcatctgaagaagctgaggaatccaccaccaacttaagaagataagctctcgaagct 321  
Db 2451 AAGCCCAAGACAGAAAGCTTGTATGACCAAGAGGCTCTGGATGAAGATATACCTCCAGTTT 2510  
QY 322 ctctatnaaaaatttaaaaglttaacaagaagcctgaatgaccttctgtgtctcgtgagacat 381  
Db 2511 CCTTTGGAGATTAATGAGCTTAATAATTAATGATGAGCTTATTTTGTGTGGTGAT 2570  
QY 382 tgcgtctctgtctaaatgaagacgttcaatcagaatattccatcaatctcttcaaaaata 441  
Db 2571 TTTGTGCAATTAATAATGAGAGCTGCAACCAAGATTTGCAATATATTTTCAAGATTA 2630  
QY 442 taacatgcttggaaataatgaataaaacacattaaag---catgatgcagaatattat 498  
Db 2631 TAGGGGTTTTACAAAGATTTATGATCTGTGAACAAAACGACAGATATGTAATTTGTTT 2690  
QY 499 ccataatctcgtcatttgggtgttaltgtctgtcagttatctcaagagatbat 558  
Db 2691 CACTATTTGTTCTAGCCCTTGGAATTTATTTCTTTATATCATTTTCTTCAGGGTTTCA 2750  
QY 559 ttacggcagaagcagggaatatttaacgatgagattaaagacgttgccttcaagcca 618  
Db 2751 CATTTGGCAAGCTGGAGAGATCTCAACCAAGCGCTCCGATACATGTTTCCCATCA 2810  
QY 619 ttttatcatcagaatattgacctgtgttgaatgaagaagaacagacagagagcttgaca 678  
Db 2811 TGCTCAGACAGAGATGTGAGTTGTTTGTATGACCTTAATAAACCAACCTGAGACATTTGACTA 2870  
QY 679 caatatgaagcagatataaagacaactcaagagagacagacaggttccagagttcgt 738  
Db 2871 CCAGGCTGCCAATGATGCTCTCAAGTTAAAGGGCTATAGTTCCAGGGTGTCTGTA 2930  
QY 739 taacacaaaatgacaactaacatggaacttcaagtatacttctcttataatgagatgg 798  
Db 2931 TTACCCCAAAATATGCAAAATTTGGACAGAAATATATATCTTCAATCATATGTTTGGC 2990  
QY 799 agatgacatctcgtacttgcgtgagtatgtctcagtaacttgcgtgacagaatgatitga 858  
Db 2991 AACATACCTTTACTTACTTACCAATTTGACCAATTCATTCAGATGACAGAGTTGTGAA 3050  
QY 859 ccgagcagaatgagatltgccaacaaagaataagcaaaacttaagaatctggaaga 918  
Db 3051 TGAATAATGTTGTGACCAACGACATGAAGAAAGAACTAGAACAGTGTGGGAGA 3110  
QY 919 taagaactgaagcttggagaataagtaactatagtgatcaatcaagaaggaagaagcct 978  
Db 3111 TCGGTACGGAAGCAATAAATACTCCGACCGTTGTTTCTTGTGACTGACGCGAAGT 3170  
QY 979 tcgagcaaatgataagaagatgctcagactcaacacagaataactcagaagaagcac 1038  
Db 3171 TTGAACATATGTATGCTAGAGTTTGACAGTACCAATACAGAAACCTTTTGAAGAAACAC 3230  
QY 1039 agatattggaagctgtatgcatcaagcaltgccttataatlttgactcaagag 1098  
Db 3231 ACATCTTTGGAATTCATTTTCCCTCACCAAGCAATGATGATTTTTCATGCTGAT 3290  
QY 1099 ggttcgaatggagccatttaataatcaagctgagaatgaacccagaagggcatgttca 1158  
Db 3291 GTTTCGGGTTTGGAGCCCTACCTTGTGGCAATTAACATCATGATGATGAGGATTTCTGT 3350  
QY 1159 taagttaacgcaatgcatatgagctatgagacatcggaagaacgctgttggctc 1218  
Db 3351 TAGATTTTTCAGAGCTGTGTTTGTGTGCTATGCGCGTGGGCAATGCTCATTTTGTCTC 3410  
QY 1219 ctgaatctcaaaagcaaatcgagggtgtgcacatgttctgtctgttggaaaagaac 1278  
Db 3411 CTGACTATGCAAAAGCAAAATATATCAGACGCCACATCATCATGTATGTAATAAACCC 3470

QY 1279 caatatagacagccgacatcaagaaggaagaaagccagacacatgtgaagaatttag 1338  
Db 3471 CTTTGATTTGACAGCTACAGCACGGAGAGCTTAATGCCAACAACATTTGAGAGAAATGTCA 3530  
QY 1339 agttcagaagctctctctctcaatcgaatgtcgcacagatgttcaatcccggtgct 1398  
Db 3531 CATTTGGTGAAGTTGTATTAACATATCCACCGACCGACATCCCATGCTTCAGGGAC 3590  
QY 1399 talccctcagatagcagcgaagaagacatgacattgtgtggagcagcagctgtgga 1458  
Db 3591 TGAGCTTGAGAGTGAAGAGGCGCACAGCTGCTGTGTTGGGACAGTGTGCTGGGA 3650  
QY 1459 aagcactctgttcaacttcttgagagacttataagcccggtcgaagaagaagtcgt 1518  
Db 3651 AGAGCACAGTGTGCTGAGCTCTGGAGCGCTTCTACAGACCTTGACAGGAAGTGTCTGC 3710  
QY 1519 ttgatgtgtgtagaataaagaatgaatgaatgacatgacatgacatgacatgacatgac 1578  
Db 3711 TTGATGCGCAAGAAATATAAGGAGCTGATGTTCAAGTGTCTCGAGCACACCTGGGATCG 3770  
QY 1579 ttccctcaagccgtgtctcttcaactcagcagatgtctgagaacatgcctatgttga 1638  
Db 3771 TGTCACAGAGCCCATCTCTTTGACTGACATGCTGAGAAACATTTGCTATGAGACA 3830  
QY 1639 aagccggtgtgtgcatatgagatcaagaagcgcgaatgacgaatcatcat 1698  
Db 3831 ACAGCGGCTGTGCTCAGAGAGAGATCTGAGGCGACCAAGAGCCACATATACATC 3890  
QY 1699 ctltatgaagctctctcagaaatacaacaagaatgttgaactgaaagagacagc 1758  
Db 3891 CCTTCATGAGTCACTGCTGCTAATAATATACACATTAAGAGTGAACAAAGAACTCAGC 3950  
QY 1759 ttctgcgcgcagaacaaagaactagctatgtcaagagctcttccaaaaaccaaa 1818  
Db 3951 TCTCTGTGGCCAGAAACAGCATTTGCCATAGCTCTGCTCTTGTAGACAGCTTATA 4010  
QY 1819 ttatatgttgatgagcacttcagccctcagataatgaagaatgagaatgtgttgc 1878  
Db 4011 TTTGCTTTTGGATTAACCCACGCTGCTGTGATGACAAAGTAAAGTTGTCACAG 4070  
QY 1879 atgcccctgataaagccgagcaggaagacatgactagtgctcactcagcagctctct 1938  
Db 4071 AAGCCCTGGACAAACAGAGAGAGCCGACCTGATTTGTGTTCTCACCGCCTGTCA 4130  
QY 1939 caatcagaagcagatltgtagtgtgtcgcacatgaaagaataagaagaagaaga 1998  
Db 4131 CCATCCAGAAATGCAAGCTTAATAGTGTGTTCAAGAAATGGCAGACTCAAGAGGATGCA 4190  
QY 1999 ctcaatcaagagctcctgagaatcgagacataatlttaagttagtgaatgacagctcag 2058  
Db 4191 CGCATCAGACGCTGCTGCGACAGAAAGCATCTATTTTCAATGTGTGAGTGTCCAGGCTG 4250

RESULT 7  
108557  
LOCUS 108557 4669 bp  
DEFINITION Sequence 3 from Patent WO 8705943. PART 02-DEC-1994  
ACCESSION 108557  
VERSION 108557.1 GI:588735  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4669)  
AUTHORS Roninson, I.B., Pastan, I.H. and Gottesman, M.M.  
TITLES COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES  
JOURNAL ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
FEATURES  
source Location/Qualifiers  
BASE COUNT 1394 a 892 c 1129 g 1254 t  
ORIGIN



RESULT 8  
LOCUS AR028671 6505 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 5 from patent US 5858744.  
ACCESSION AR028671  
VERSION AR028671.1 GI:5940644  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 6505)  
AUTHORS Baum, C., Stocking-Harbers, C. and Osterlag, W.  
TITLE Retrieval vector hybrids and the use thereof for gene transfer  
JOURNAL Patent: US 5858744-A 5 12-JAN-1999;  
FEATURES  
source 1. 6505  
/organism="unknown"  
BASE COUNT 1719 a 1506 c 1606 g 1674 t  
ORIGIN  
Query Match 34.5%; Score 712.4; DB 6; Length 6505;  
Best Local Similarity 60.1%; Pred. No. 5.4e-141;  
Matches 1262; Conservative 0; Mismatches 796; Indels 42; Gaps 3;  
OY 1 cgaagcaaggtcgcgaactcaatcgtgtgtagcacaccgaacttctactatctgaagtcag 60  
DB 3543 CCAGAAAAGTCGCGCCACCATTTGTGATAGCTCATCTTTGCTACAGTTCTTAATGCTG 3602  
OY 61 attgaattgtagcccttaagaatgtagtgcgtaggaagaagaagacatctgaactca 120  
DB 3603 ACGTCATGCGTGGTTGTCATGATGAGCATTTGTGAGAAAAGAAATATATGTAAGTCA 3662  
OY 121 tggcaacaagagtgctatattatctcctgtgtagtcacagatataaaaaagctgag 180  
DB 3663 TGAAGAAGAAAGCATTTACTTCAAACTTGTCACAAATGCAGACAGCAAGAAATGAAGTTG 3722  
OY 181 aacagatgagtcacatgacatattcta-----ctgaagaagaagccaa 223  
DB 3723 AATTAGAAAATGCAAGCTGATGATCAAAAGTGAATTTGATGCTTGGAAATGTCCTTCA 3782  
OY 224 ctccactctctgcacgtctgtagaagacatcaag-----ctgaagaagaagccaa 261  
DB 3783 ATGATTCAGATCCAGTCTTAATTAAGAAAAGATCACTCGTAGAGTCCGTGATGAC 3842  
OY 262 actcattgacaaggtcgtgaggaatccaccaatcaagaagataagctctcctgaagct 321  
DB 3843 AAGCCCAAGACGAAGCTTAGTACCAAGAGGCTCTGATGAAAGTATACCTCCAGTTT 3902  
OY 322 ctctatataaaatttaagttaaacaagcctgaagccttctgtgtcttggtggagcat 381  
DB 3903 CCTTTTGAGATATGTAAGCTTAATTTACTGATGATGCTTATTTTGTGTTGGTGTAT 3962  
OY 382 tggcttcgtcttaaatgaaactgttcacagtatttccatcattccatcttgcataaata 441  
DB 3963 TTTTGCCATTATTAATATGAGAGCCTGCAACAGCATTTGCAATATATTTTCAAAAGATTA 4022  
OY 442 taacatctgttggaaataatgataaaacacatlaaag---catgatcagaatattat 498  
DB 4023 TAGGGGTTTTCACAAAGATGATGATCTGAAACAAAAGCAAGATAGTAAGTGTGTTT 4082  
OY 499 ccaatgatactcgtaatttgggtgttattctgtccttgcagtattatcagaggagattat 558  
DB 4083 CACATTTGTTCTAGCCCTTGGAATTAATTTCTTTATTAATTTTCCCTCAGGGGTTTCA 4142  
OY 559 tttaagcagagcaggggaaattttaacgtagagattaagaacactgctctcaagaagca 618  
DB 4143 CATTTGGCAAAAGTCGAGAGATCTCCACCAAGGGCTCGATACATGTTGTTTCGATCA 4202  
OY 619 tgtatatacagagatactcctgtgtttagtgaagaagaagaacagcagagagcttgaca 678  
DB 4203 TGCTCAGACAGAGATGTGAGTTGTTGATGACCCCTAAAAAACACCATGAGAGATTGACTA 4262

OY 679 caatatagccatagatatagcacaatccaagaagcaacaggttccagattgctgct 738  
DB 4263 CAGGCTGCGCAATATGATCTGCTCAAGTTAAAGGGGCTATAGTTCCAGGCTGCTGTA 4322  
OY 739 taacacaaatgcaactacatctggaacttcagttatcatctcttatatgagatgg 798  
DB 4323 TTACCAGAAATATAGCAATCTTGGACAGAGATATATATATCTTCAATCTATGTTG 4382  
OY 799 agatgacatctcgtattctgagtagtctccagtaacttgcgcgtgacgaatgattgaa 858  
DB 4383 AACTAACACTGTTACTCTTAGCAATTTGTAACCATCATGATTCAGAGAGAGTTGTA 4442  
OY 859 ccgcagcaatgactgattgctccacaagaatgaagaagaacttaagatctgtgaaaga 918  
DB 4443 TGAATAATGTTGCTGGAAGAGCATGAAAGATMAAAAGAACTAGAAAGCTGCTGGAGA 4502  
OY 919 tagcaactgaaagccttggagaaatatacgtactagtgctcatatacgaaggaaagcct 978  
DB 4503 TCGCTACTGAAACATAGAAAACCTCCAGACCGTTGTTCTTTGACTCAGAGAGCAAGT 4562  
OY 979 tcgaagcaatgtagaagaatgcttcagactcaacacagaataacctcgaagaagcac 1038  
DB 4563 TTGAACATATGATGCTCAGAGTTTGAGGTACATACAGAAACTCTTGAGAGAAAGCAC 4622  
OY 1039 agatattggaagcgtgtatgcatcagcgaatgcttataatttgcctatgcaagca 1098  
DB 4623 ACATCTTTGGAAATTCATATTTCTTCCACCCAGGCATATATATTTTCCATATGCTGAT 4682  
OY 1099 gtttcgatttggagaccatatttaactcaagcgtgacgaagtaccacgaaggacatgta 1158  
DB 4683 GTTTCGGTTTGGAGCTCTACTTGTGTGACATTAACCTATGATAGCTTTAAGATGTTCT 4742  
OY 1159 taatttactgcaatgtgcatatgtagtctgacatgacatgcgaagaacagctcgtttgctc 1218  
DB 4743 TAGATTTTCAAGCTGTTGCTTGTGTCATGCGCCGTGGGCAAGTCAATTCATTTGCTC 4802  
OY 1219 ctgaatattccaagccaaatccggtggtgcacatctgttgccttgttggaaagaagac 1278  
DB 4803 CTGACTATGCAAAAGCCAAATATATCAGCAGCCCAATCATATGATTCATTTGAAACCC 4862  
OY 1279 caaatatagacagccgagctcaagaaggaagaacacagacatgtagaagaattag 1338  
DB 4863 CTTTATTTGACAGCTACAGCAGAGGCTTAATGCTCCAAACATTTGGAAGAAATGCA 4922  
OY 1339 agttcgaagagctccttcttctatccatgctgcagatgttttcaatccctcgtgct 1398  
DB 4923 CATTTGGTAGAGTTGATTTCAACATATCCACCCAGCCGACATCCAGTGTTCAGGAC 4982  
OY 1399 tatccctgaattgagcgaaggaagacagatagcaattgtgtggagcagcggtgtgga 1458  
DB 4983 TGAGCCTGAGAGTGAAGAGGCGCAGAGCGTGGCTCTGTGGGCGAGCATGTGCTGGGA 5042  
OY 1459 aagacactctgttcaactctgcagagacttttagaccccgtagaagaagaagtgct 1518  
DB 5043 AGAGCAGAGTGTCCAGCTCTGAGCGGTTCTAGACCCCTTGCCAGAGAAAGTGCTGC 5102  
OY 1519 ttgattgtgtgtagtcaaaagaatgaatgtagtcaatgtgctcgttcccaaatagcaatgc 1578  
DB 5103 TTGATGGCAAGAAATATAAGGAGCATGATGTTCAATGTGCTCGACACACCTGGGCATCG 5162  
OY 1579 ttctcgaagcctgtgctctcctcaactgcagcaatgtctgagaacatgcctatggtgaca 1638  
DB 5163 TGTCCAGAGAGCCCATCTCTGTTTGAATCTCAGCATTTGCTGAGAACATCTCCATATGAGACA 5222  
OY 1639 aagccgctgtgtgcatatgagatgagatcaagaagccgcaatgcaaatatccat 1698  
DB 5223 ACAGCCGGGTGTGTCAAGAGAGATGCTGAGGGCAGCAAGAGGCAACATATACATG 5282  
OY 1699 ctatttgaaggtctcccttagaataacaacaagattgtagtgaaggaagcagc 1758  
DB 5283 CCTTCATGAGTCACTGCTCTAATAATATATAGCACTAAAGTATGAGAGCAAAAGAACTCAGC 5342

OY	1759	ttcttcgcggcccaaaacaaagactagctcatcttgcaagggtctcttcctcaaaaccccaaa	1818
Db	5343	TCTGTGTGGCCAGAAACACGCGATTGGCCATAGCTCTGTGCTCCCTTTGTAGACAGCTCAT	5402
OY	1819	tttatctgttgagatgagccacttcagccctctgataatgacagatgagaagtglttaagc	1878
Db	5403	TTTTGCTTTTGGAGGAAGCCACGCTCGCTGTGATTCAGAAAGTAAAGGTTGTCAG	5462
OY	1879	atgcgccttgataaagccaggaacggggaaggaacatgcccctagtgctactcaagcctcttg	1938
Db	5463	AAGCCCTGGAGAAAGCCAGAGAGGCCGCGACCTCGCTTGTGTTGCTTCACCCGCTGTCCA	5522
OY	1939	caattcgaacgcagatttgatagatggtctctgcacaaatggaagaataaagagaagaagaa	1998
Db	5523	CCATCCAGAAATGCAGACTTATATAGTGTGTTTCAGATGCGAGACTCAGAGGAGCATGTGCA	5582
OY	1999	ctctacaaagatcctcttaagaatcgcgacatataatttaagttagtgatgaatgaagttag	2058
Db	5583	CGCATACGACGCTCTGTGGCAGAGAAAGCATCTATTTTCAATGTGCTCATGCTGCAGGCTG	5642

	RESULT	9					
	AR028672						
	LOCUS	AR028672	9318 bp	DNA	PAT		29-SEP-1999
	DEFINITION	Sequence 6 from patent US 5858744.					
	ACCESSION	AR028672					
	VERSION	AR028672.1	GI:5940645				
	KEYWORDS	.					
	SOURCE	Unknown.					
	ORGANISM	Unknown.					
	REFERENCE	Unclassified.					
	AUTHORS	1 {bases 1 to 9318)					
	TITLE	Baum,C., Stocking-Habbers,C. and Osterlay,W,					
	JOURNAL	Retroviral vector hybrids and the use thereof for gene transfer					
	FEATURES	Patent: US 5858744-A 6 12-JUN-1999;					
		Location/Qualifiers					
	source	1..9318					
		/organism="unknown"					
	BASE COUNT	2377 a 2217 c 2364 g 2358 t					
	ORIGIN						
		2 others					

Query Match	34.58;	Score 712.4;	DB 6;	Length 9318;
Best Local Similarity	60.18;	Pred. No. 5.4e-141;		
Matches 1262; Conservative	0;	Mismatches 796;	Indels 42;	Gaps 3;

[illegible]

QY	382	tgagcttcgttctaaatggaagcgtctacacagatattcttcacactcttgcgaataat	444
QY	3922	tttgctccattatTAANAGGAGGCCCTCGAACACGACATTTCGAAATATTTTCMAAGATTA	3981
QY	442	taaccatgltttggaataatgataaaccacatlaag---catgatgcagaatattat	498
Db	3982	TAGGGGTTTTACAGAAATGATGATCCTCGAAMAAACAGACAGATAGTAACCTGT	4041
QY	499	ccatgatattgcgaattcttggtgtatttgcgtctgcagtattcttcagcaggaatat	558
Db	4042	CACATTTGTTCTAGCCCTTGGGAATTAATTTCTTTATATCAATTTTCCCTCAGGGTTCA	4101
QY	559	tttacgagcagcagagggaatttlaacgatgagatlaagacacttggccctcaagcca	618
Db	4102	CATTGGCAAGCGTGAGAAATCTCCACCAAGCGCTCGAATACATGTTTTCCGATCA	4161
QY	619	tgttatatacggatattgccttglttttgatgaaagaagaaacagcagcagaagcttgaac	678
Db	4162	TGCTAGACAGAGATGATGTTGGTTTGAATACCCCTTAANAACCACTGAGCATTGGACTA	4221

QY	679	caatlttagcctatagatatagcaacaattcaaggagcaacaggttccagagatttgcgtc	738
Db	4222	CGAGGCTGGCCAAAGATGATGCTGCTCAAGTTAAAGGGCTATAGATTCCAGGCTTGCTGTAA	4281
QY	739	taacacaanaatgcactaacaataggagcttcagltatcatcttccttlatatatgata	798
Db	4282	TTACCCAAATATATGCAAAACTTGGGACAGAGATATATTTATCTTCATCATGTGTTGGC	4341
QY	799	agatgaccttcctgattcttcgaatattgcttcagtaatttcggtgcgaagaatagattgaa	858
Db	4342	AACCTACACTTTACTCTTAGCAATTTGTACCCATCTTCATAGCAAGAGTTGTTGAAA	4401
QY	859	ccgcgcacatctactcgttatcttgcacacaagaatagaagaagacttaagcatcttgcgaa	918
Db	4402	TGAAATATGTTGTCGACACAGCACTGTAAAGATATAGAAAGACATAGAAAGTGCTGGGAGAA	4461
QY	919	tagcaactcgaagctctttggagaatatagctatctatgtgtcattaaacaaggaaaaagct	978
Db	4462	TTCGTACTGGAAGCAATATGAAAACTTCGGAACCGCTGTCTTTCTTTGATCATGGAGCGAAGAT	4521

QY	979	tcgagcaaatgtatagaagagatgcttctaagaactaaacacgaataactctbgaagaagac	1038
Db	4522	TTGAACAAATATATATGCTCAGAGTGTGGAGGTACATACGAAACTCTTTGAGGAAAGAC	4561
QY	1039	agattattggaagctgttatgcatctgaagcagcatgccccttatataatttgcctcagcag	1098
Db	4582	ACATCTTTGGAAATTAACATTTTCCCTTCCACCCAGGCATGATGTTTTCCTATGCTGGAT	4641
QY	1099	gatttcgaattggagcctalttaattcaagcttgaacgaatgaacccagaggacatgtca	1158
Db	4642	GTTTCCGGTTTGGAGCCACTGTGTGGCATATAACTCATGACCTTTGAGATGTTCTGT	4701
QY	1159	tagtltttactgtcaattgcatatagagctatbgcatcgtgaataaacgctcgttttgctc	1218
Db	4702	TAGATTTTTCAGCTGTTGTCTTTGGTCCATGGCCGCGGGCAGACCACTTCATTTTGGTC	4761
QY	1219	ctgaatattccaagaacccaatccggggctbgcgatctgttgcctgttggaaagaagac	1278
Db	4762	CTGACTATAGCCAAAGCCCAAAATATACGACGCCACATCATGATCATTTGAAAAACC	4821
QY	1279	caaatatagacagccgcaqtcaagaagaagaaaaagccagacacatgttgaaggaaattag	1338
Db	4822	CTTTCATTGGACAGCTACAGCACGGAAGGCTTAATGGCGAACAACATGTGAAGAAATGTCA	4881
QY	1339	agttccgagaagctcttcttctctatccatgtgtcccgagatgttttcaaccctccgtgct	1398
Db	4882	CATTTGTGGAAGTTGATATTAACATATCCACCCGACCGGACATCCCATGGTCTTCAAGGAC	4941
QY	1399	tatccctcagatattgagcaggaagaagccagtagcatittgtgggaagcaagcgctgtgga	1458
Db	4942	TGAGCCTTGAGAGTGAAAGAAGGACCAGACGCTGGCTCTGGTGGGACAGATGGCTGTGGGA	5001





```

Db      2933  GTTCCGGTTGGAGCCTTACTGTGGCCACATTAACATCATGAGCTTGTGAGGAGTTCGT 2992
Oy      1159  tagttttactgcaatgatgatgagctatggtccatcggaaaaagctggtttgctc 1218
Db      2993  TAGATTATTCAGCTGTTGCTCTTGGTCCATGGCCGTGGGCAAGTACATTTCTC 3052
Oy      1219  ctgatatccaaagccaaatcggggctgcacatcgtttctgcttctgagaaagaac 1278
Db      3053  CTGACTTGGCAAAAGCCAAATATCAGCAGCCCATCATCATGATCATTTGAAAAACC 3112
Oy      1279  caaatatagacagccgagtcgaagaagaaagaaagcacaatggaaggaattag 1338
Db      3113  CTTTGTATGACAGCTACAGCAGCAGGAGCCTTAATGCCGACACATTTGGAAGAAATGCA 3172
Oy      1339  agtttgaagagctcttcttctatccatgctgcggccagatggtttcatccctcggtc 1398
Db      3173  CATTGTGTGAAGTGTATTAACATATCCACCCAGCCGAGCATCCCATGCTTCAAGGAC 3232
Oy      1399  taccctcagatattgagcaggaagaagacagtagcatctgtgtggagcagcgctgtgga 1458
Db      3233  TGACCTCTGGAGGTGAAGAAGGCGCAGACGCTGCTGCTGGTGGCAGCAGTGGCTGGGA 3292
Oy      1459  aaagcactctgttcaactctgcagagactttagaccctgtgcaaggaagtagtctgt 1518
Db      3293  AGACACAGTGTGTCCAGCTCCTGAGCGGTTCTAGCAGACCCCTTGCAAGGGAAGTGTGC 3352
Oy      1519  ttgtgtgtgtgtagcacaagaatattgaatgtagcagtgctcgtttcccaatagcactg 1578
Db      3353  TTTGTGGCAAAATAAATAAGCGACTGAATGTTCAGTGGCTCCAGCAGCAGCTGGGATCG 3412
Oy      1579  ttctcagaagcgtgtgctcttcaactctgcaagcattgtctgaaacatcgctatgtgaca 1638
Db      3413  TGTCOCAGAGCCCATCTCTGTTGATGCTGACCATTTGCTGGAACAATTCCTATGAGACA 3472
Oy      1639  acagcgtgtgtgtgtagatgtagatcaagaagccgcaatgcaaatatccatt 1698
Db      3473  ACAGCGGGGTGTGTACAGAGAAGATTTGAGGGCAGCAAGAGGCGCAACATACATG 3532
Oy      1699  ctttattgaaggtctccctctgagaatacaacaagaattggaagtcgaagcagc 1758
Db      3533  CCTTATCATGAGTACGCTGCTTAAATATATACACTAAAGTACAGACAAAGAACTCAGC 3592
Oy      1759  ttctgtgcgcgcgaagaacaagaactagctattgcaagggctctctccaaaaaccataa 1818
Db      3593  TCTCTGTTGGTGGAGAAACAAGCATTTGCCATAGCTGCTGCTTGTAGACAGCTCTATA 3652
Oy      1819  ttctattgtgtgtagagccactttagccctcgataatagtagcagtagaaggtgtctagc 1878
Db      3653  TTTTGTCTTTTGGATGAAACCAAGCTGCTGATACAGAAAGTGAAGGTTGTCCAG 3712
Oy      1879  atgcctctgataaagcaggaagcaggaagcagctgctgtagtgcactcaagcctctgt 1938
Db      3713  AAGCCCTGGACAAAGCCAGAGAAGGCCGACCTGCTGATGTGATTTCTCAACCCTGTCA 3772
Oy      1939  caattcaagaagcagcagattgtgtagtctgtcgcacaaatggaagaataagaagaagaa 1998
Db      3773  CCATCCAGATGCAAGACTTAATAGTGTGTTCAGAAATGGCAGAGTCAAGAGCATGCA 3832
Oy      1999  ctcatcaagagctcctgagaagaatcgagacatatatttaagttagtgaagtagcagctc 2058
Db      3833  GCATTCAGAGCTGCTGGCAGAGAAAGGATCTATTTTTCAATGGTGTGATGTCCAGGCTG 3892

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 1 to 4186)
AUTHORS
Stocker,P.J., Steinel-Crespi,D.T. and Crespi,C.L.
TITLE
P-glycoproteins from macaca fascicularis and uses thereof
JOURNAL
Patent: WO 0123565-A 1 05-Apr-2001.
GENEST
CORPORATION (US)
FEATURES
Location/Qualifiers
1..4186
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
100..3942
/codon_start=1
/feature="unnamed protein product"
/protein_id="CAC37764.1"
/db_xref="GI:13923887"
/translation="MDLEDDRNCGAEKKNFKLNKSKKDKKERKPPVSVFSPFRYSN
WLDKLMVYGTLLAIIHAGCLPLMLVDPDMTDPANAGNLDGLALFNSSNITPTV
PVNKLDEMRITAVITSGIAGVLAAYIQVSPWCLAAROIHKIKRPFHIMROI
GMPDVHDEGLNTRLDVSKINEIGDKIGMFOSMAFFFGFIVGTRGKRLTIVT
LAISPVGLSAVAWAKILSFDDKELLAAGAAVEVLAALRTVIAFGGKLEDER
YNNLEBAKIGIKKATITANISGAFLIIVASALAFVGTTLVLSKEYSISQVLT
FSVLIGAPSVGOASPIEAFANARGAPEIFKIIDNKPISIDSKSGKHPDKGNL
EPRNVPSPSKREVKILKGINLKXGSGTVALVGGSGKSTTVOLMRIDPTEGM
VSDGODIRITNRLREIIGVYSQEPVLFATITANIRGRREDTMRERAKRAN
AYDFIKLPQKFTLVGERAQLSGQKQKQIATARLVNRPKLLIDEATSLDITSE
AAVQVALDKARKGRITVIAHRLSTVBNADHVFDDGVIVEGNDELKKEGILFK
LTMOTNAGIELENADESKSEIDPLEMSHDVSGSLIRKSTRSVSGSOQDKRL
STKEALDESPVSPVWRKMLNTEMPVFGVCAIINGLOPAVAFVSKILGIFT
RNDWAEKRONSNLFSLLPLVIGITVSEIFPEFGFGPRGAKVITKRLRMVFRSMR
QDVSWEDPKNNTGALTURLANAAVKKAGISRIALITONIANLGTITISLYGMO
LTLTLAIVPIIAIAGVENKMLSGALKKKLEBAGKATATAIENPRVVSLEBO
KLFHMVDOSLOVYRNSLRKHLFGITFSTQAMVFSYACRFAYLVAHLSMSE
DVLIVSAVFGAMAVGOVSPFADYAKAKAVSAHIIIMIEKPLIDYSTELKENT
LEGNVNEVEVNPRLDIPVLOGSLKVKKQTLALVSSGCKSTVOLLEERYD
PLAGVILDKKEIKOLNWOVLRAHLGIVSQEPLIPCSISENIAVGDNSRVSEEV
RAKKEANTHAFISLPKYSSTRYGDGKTGSGGOKRIATRLABVQPHILDEATS
ALDSEKVVQEAALDKRBSRGTCTIVIAHRLSTIONADLIYVONGVKEKHTMOQLLA
OKIGTSMVSVQAGAKRO"
BASE COUNT 1226 a 799 c 1039 g 1122 t
ORIGIN
Query Match 34.4% Score 710.8; DB 6; Length 4186;
Best Local Similarity 60.0%; Pred. No. 1.2e-140;
Matches 1261; Conservative 0; Mismatches 797; Indels 42; Gaps 3;
Oy      1  cgagaagaagtcgagactcaatcgtgtgtagcacacagcttctactatctgaagtcaag 60
Db      1826  CCAGAAAGGTGGGACACCATTTGTGATAGCTCATGCTTTGTCTAGGGTGTGTAATCGG 1885
Oy      61  attgatctgacacctaaagaatggaatgctgtgcggaagaagacagacatgctgaactaa 120
Db      1886  ACGTATGCTGTGTTGCATGATGAGTCATTTGTGTGAGAAAGAAATGATGATGAGCTCA 1945
Oy      121  tggcaaaagaggtctatattactactgt-----gatgcaagga 163
Db      1946  TGAAGAGAAAGGCACTTACTTCAAACTTGTCAACATCCAGACAGCAGGAATGAAATTG 2005
Oy      164  tattaataaagcgtgtgaacagatgtagtcaatctactactatctgaagaagaagacca 223
Db      2006  AATTGAAATAATCCAGCTGATGAAATCCAAAGTGAATGATTACTTGGAAATGTCTTAC 2065
Oy      224  ctcaactctctgcaactctgtgaagagcataag-----tcag 261
Db      2066  ATGATTTACAGATCCAGTCTTAATAAGAAAAAGATCCACTCGTAGAGTGTCGGTGAATC 2125
Oy      262  attcatctgacaagcgtgaggaatccacccaactcaagaagataagctctctgaagctct 321
Db      2126  AAGGCCAAGACAGAAAGCTTACTACCAAAAGAGGCTGTGATGAAGTAACTTCCAGTTT 2185
Oy      322  cctataataaattttaagttaaacaagcctgaaatgtagccttctgtgtctcggggaacat 381

```





Medicine, Stanford, CA 94306, USA

FEATURES

Source

Location/Qualifiers

1. .4192

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="7"

/map="7q21.1"

/cell\_type="drug-resistant uterine sarcoma"

/cell\_line="MES-SA/DxP"

/clone="8.1"

1. .4192

/gene="mdr1"

/note="multidrug-resistance gene"

120. .3959

/function="transporter protein"

/note="multidrug-resistance"

/codon\_start=1

/product="p-glycoprotein"

/protein\_id="AAB69423.1"

/db\_xref="gi:2353264"

/translation="MDLEGDRNGAKKKKRNPKLNKSKKKKKPKPYVSVSMRYSN  
WLDKLYNVGFLAIIHGAGLPLMLVFGEMTDIFPANGNLEDMSNTNPSDINDTS  
FPMLEDMTRYAYYSIGAGVLAAYIOVSFCLAGRQIHKIRKOFPAIRKOEI  
GMPFVHDVGEINLTLDVSKINSGIGDKIMFQSMATPFTGFIYVGRMKTLVI  
LAISPVGLSAAMAKILISFTDKELLAYAGAAVEVLAIRTVIAFGGCKELER  
YKNLEAKRIGIKKAITANTISGAFLIYAVALAWGTTIVLGEYSIGQVLY  
FSVLGASVGSQASPIEAFANAGAAETIKITDKNSIDSYSKSGKPRNIGNLE  
FRNVHSTPSRKVKYILKGLNKVOSQTVLVGSGGKSTVQLMRLDPYEGAV  
SVGODIETINVRFLRETIIVVSOEPEVLAFTIENIRYGENVTMDRIEAEVANA  
YDFIMKLPKHKEDTIVGERGAQISGOKORIAIARALVNPRIILLDETSALDESEA  
VVOVALDKARKGRTTIVAHSLTVRNADVIAGDDGIVYKGNHDELMKGIYFKL  
VTMOTAGNEVELENADESKSEIDALEMSSNDSLSLRKSTSRSVSGAORRKS  
TKELADESIPVSPFMRIMKLNTEMPYVNVCEALINGCOPARATFEKTIYFPR  
IDDPETKRQNSNLSLFLALGITSFTFFLQFTFGAGELITRLRYVFRSLRQ  
DVSWEDPKNTGALTTRLANDAAOVKAISRLAVITQNTANGTGLIIFIGWOL  
TLLLAIVPIIAIGVEMKMSGALDKKELGAGKIAEATENFTVSLIOEOK  
FEHMAOSLOVYRNSLRKAHIFGTFSEFTQAMVYFVAGCFRGAYVAHKMSFED  
VLIVFSAVFGAMAVGOVSPAPYAKKISAHHIMIEKTPILDSYSTGIMPNIL  
EGNVTGFEVYVNPTRPDIPYLOGISLEVKKGQTLALVSSGCKSTVYOLLERYDP  
LAGKVLDEKIKRLNVOMLRHGLGISQEPILDCSIAENIAYGDNRSVVSQEIYR  
AAKANIHAFIESLPNKYSTVVGKGTOLSGGOKORIAIARALVNPRIILLDETS  
LDTSEKVVQALDKARKGRTCIYIAHRLSTIQNALDILVFPNGRVKKEHGHQDLAQ  
KGIYFSMVSVOAGKRRQ"

659

/variation

/gene="mdr1"

/citation=[1]

/replace="c"

673. .674

/variation

/gene="mdr1"

/citation=[1]

/replace="tt"

1351

/variation

/gene="mdr1"

/citation=[1]

/replace="c"

1428-1429

/variation

/gene="mdr1"

/citation=[1]

/replace="ttc"

3551

/variation

/gene="mdr1"

/citation=[1]

/replace="c"

BASE COUNT 1243 a 797 c 1028 g 1124 t

ORIGIN

Query Match 34.4%; Score 710.8; DB 9; Length 4192;  
Best Local Similarity 60.0%; Pred. No. 1.2e-140;  
Matches 1261; Conservative 0; Mismatches 797; Indels 42; Gaps 3;

QY 1 cgagcaaggtcgactacaatcgtagcacacgacttctactatctgaagtgcag 60

Db 1843 CCAGAAAGGTCGGACCAACATTTGTAGTACGTCATCGTTTGCTACAGTGTGAATGCTG 1902

QY 61 attgatgtgtaacctaaagatgtaagtctggtcggaagaagggacaatctgaactaa 120

Db 1903 ACGTATGCTGTGGTTTCCATGTGAGTCATTTGTGGAAGAAATCATGTAGTCA 1962

QY 121 tggcaaacgagctctatattcttactctgtaagtctcaagatataaataagctgag 180

Db 1963 TGAAGAGAAAGCCATTACTTCAAACTTGTCCACATGCACAGCAGAGAAATGAGTTG 2022

QY 181 aacagatgagtaactgacatattcta-----ctgaagaagaaccaca 223

Db 2023 AATTGAATAATGCACCTGATGATCCAAAGGAAATTTGATGCTTGGAAATGTCTCAA 2082

QY 224 ctcaactctctgcaactctgtaagatcaag-----taag 261

Db 2083 ATGATTCAAAGATCCAGTAAATAGAAAAGATCAACTCGTAGAGTGTCCGTGATCAG 2142

QY 262 acttattgacaagctgaggaatccaccacaatctaagaagataagcttctgaagctc 321

Db 2143 AAGCCCAAGACGAAAGCTTACTACCAAGAGGCTCTGATGANAAGTATACCTCAGTTT 2202

QY 322 ctctaataaaatttaaaagttaacaagcctgaatggccttggctctcgaggacat 381

Db 2203 CCTTTGGAGATTTAGAAAGCTAAATTAACGATAGGCTTAATTTGTTGTTGTTAT 2262

QY 382 tggcttcgtcttaaaagaaactgtctatccaatatttccatctcttgaaaaaata 441

Db 2263 TTTGGCCATTATTAATGAGAGGCTGCACACGCACTTTGCAATATVATTTTCAAGATTA 2322

QY 442 taaccatgtttgaaataatgataaacaacataaag----catatgcagaatattc 498

Db 2323 TAGGGGTTTTCACGAATTTGATGATCCTGGAACAAAGCAGATAGTAACTGTTT 2382

QY 499 ccatgatalctgcatcttgggtgtatcttgccttgcagtattcatlgaaggatatt 558

Db 2383 CACTATTGTTCTTACCCCTTGGAATTAATTTCTTTATTAACATTTTCCCTCAGGTTTCA 2442

QY 559 tttaagcgacagcaggggaaattttaaagatgagattaaagaaactgtgcttcaagcca 618

Db 2443 CATTTGGCAAACTGGAAGATCTCCACCAACGGGCTCCGATACATVATGTTTCCATCA 2502

QY 619 tgttatcagaatattgcctgttctgtaagaagaagaacagcagagagctgtgaca 678

Db 2503 TCCCTCAGACAGATGTGAGTTGTTGATGACCTTAATAAACCCCTGAGAGATTGACTA 2562

QY 679 caaattagccataagataagacaacaattcaagagcaacaggttccaagattgagcgtct 738

Db 2563 CCAGGCTGCGCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGTGTA 2622

QY 738 taacaacaatgcaactaacaatggaacttccagtatacttcccttatataatgagtg 798

Db 2623 TTACCCAGAAATATACCAATCTTGGACAGGAATATATATATATCTCATATGTTGGT 2682

QY 799 agatgacatctctgattctcagatctgtccagtaacttgcgtgacagaaatgattgaa 858

Db 2683 AACTTAACACTGTACTCTTAGCAATTTGTAACCATCATTTGCAATATACAGAGAGTTGTA 2742

QY 859 ccgacgaatgactgattgtccaaacaagaataagcaagaacttaagaatgctgaaaga 918

Db 2743 TGAATAATGTTGCTGACACAGCATGAAAGATTAAGAAAGAACTAGAAAGTCTGGAGA 2802

QY 919 tagcaactgaagcttctggaagaataagcttactatagctgtctatacaagaaggaaagc 978

Db 2803 TCGCTACTGAAACCAATAGAAAACCTTCGAAACGGTTGTTTGTGACTGACGAGCAAGT 2862

QY 979 tggagcaaatgtaataagaagatgctgacacccaacagaagaatacctcgaaagaagcac 1038

Db 2863 TTGAACATATGTATGCTCAGAGTTTGCAAGTACCTACGAAACCTTTTGAAGAAAGCAC 2922

QY 1039 agattatggaagcgttatgcatcagccatgcttataatatttgcattatgagcag 1098

Dn	2923	ACATCTTGGATTACATATTTCCCTTCACCCAGCAATGATGATTTTCCATGCTGAT	2982
Oy	1099	ggttcgatttggaagccatttaataacagctgagcaaatgaccccaaggcattgtca	1158
Dn	2993	GTTTCCGGTTTGGACCTACTTGGTGGACATAACTCATGAGCTTTGAGGATGTTCTGT	3042
Oy	1159	tagttttacttgcattcatalatgagctatgagccatgcgaaanaacgctcgtttgac	1218
Dn	3043	TAGATTTTTCAGCTGTTCTTGGTGGCATGCGCGTGGGGAACATTCATTTTGGCTC	3102
Oy	1219	ctgaatattccaagaacaaatcgagggtgcgcactcgtttgcttctgttgaaanaaac	1278
Dn	3103	CTGACTATGACCAAAAGCCAAAATATATCATCATCATCATCATCATCATCATCATCAT	3162
Oy	1279	caaatatgacacccgcgcgtcaagaagggaanaaacacacatgtaaggaattag	1338
Dn	3163	CTTGTATTTGACAGCTACAGCAGGAGGCTTAATGCTTGAAGGAAATGTC	3222
Oy	1339	agtttcgagaagctccttcttctatccatcgtccagatgtttcattcctcgtgct	1398
Dn	3223	CATTTGGTGAAGTTGATTTCAACTATCCACCCGACGACATCCCATGCTTCAAGGAC	3282
Oy	1399	latccctaatatgagcgaggaagaagatgacatttctgtgggagcagcggtgtgga	1458
Dn	3283	TGAGCTTGAGGTGAAAGGAGGCGCAGACGCTGCTGCTGCTGGGAGCAGTGGCTGGGA	3342
Oy	1459	aaagactctcgttcaactcttgcagagactttatgaccccggtgcagaagacaagtctgt	1518
Dn	3343	AGAGCAGAGTGGTCCAGTCTCTGAGGCGGTTCTACAGACCCCTTGGCAGGAAAGTGTGC	3402
Oy	1519	ttgatggtgtgatgacaaagaatgatgatgacgtgctccgttcccaatagacatcg	1578
Dn	3403	TTGATGGCAAGAAATTAAGCAGCATGATGTTCACTGCTCCGACACACCTGGGCATCG	3462
Oy	1579	ttccctaaagacgctgtgctcttcaactgcagcaatgtctgcagacaatcgctatgga	1638
Dn	3463	TGTCAGAGAGGCCATCTCTGTTGACTGACAGATGCTGAGAAACATTTGGATGAGAGA	3522
Oy	1639	acagccggtgtgtgcattatgacatcaagaagccgcaaatgacgaataatcatt	1698
Dn	3523	ACAGCCGGGTGGTGCACAGGAAAGATTTGAGGGCAGCAAGAGGCAACATCATATG	3582
Oy	1699	ctttatgagaagctctcctctgagaataacaacaagaagtgtgactgaaaggacagc	1758
Dn	3583	CTTTCATGAGTCACTGCTCTAATTAATATATGACATTAAGTATGAGACAAAGGAATCTGAC	3642
Oy	1759	ttctctgagcgagcaaaacaaagactagctatgtcaagggtcttctccaaaaccacaa	1818
Dn	3643	TCTGTGTTGGGAGAAACAGCATTTGCCATATGCTGCTGCTTGTAGACGCTCATTA	3702
Oy	1819	ttttatgttggatgagagcacttcaagccctcgataatgacagtgaagaagtgtgtcagc	1878
Dn	3703	TTTTGCTTTTGGATGAAACACAGCTGCTGATATGAGAAAGTAAAGGTTGTCCAAAG	3762
Oy	1879	atgacctgtgataaagcgagcgaggaagatgctctagtgttcaatcagagcctctg	1938
Dn	3763	AAGCCCTGACAAAGCCAGAGAAAGGCCACCTGCTATGTGTGCTCAACGCCCTGTCCA	3822
Oy	1939	caatcagaacgagcagattgtatgtgttctgcacaaatgaaagaataaaggacaagaa	1998
Dn	3823	CCATTCAGAAATGACACTTAATAGTGTGTTTCAAGATGAGCAAGTCAAGGAGCATGGA	3882
Oy	1999	ctcactcaagagctcctgtgagaatcgagacatatatttaagttaagtgaatgacagtcag	2058
Dn	3883	CGCATCACAGCTGCTGACAGAAAGGCAATCTATTTTCAATGTGTGCTCCAGGCTG	3942

RESULT 13  
AX108656 4195 bp DNA PAT 30-Apr-2001  
DEFINITION Sequence 3 from Patent W00123565.  
ACCESSION AX108656  
VERSION AX108656.1 GI:13923888

KEYWORDS	crab-eating macaque.
SOURCE	Macaca fascicularis
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca.
REFERENCE	1 (bases 1 to 4195)
AUTHORS	Stocker, P.J., Steinmetz, D.T. and Crespi, C.L.
TITLE	P-glycoproteins from macaca fascicularis and uses thereof
JOURNAL	Patent: WO 0123565-A 3 05-Apr-2001;
GENTEST	CORPORATION (US)
FEATURES	Location/Qualifiers
source	1..4195
CDS	/organism="Macaca fascicularis" /db_xref="taxon:9541" 100..3951 /note="unnamed protein product" /codon_start=1 /protein_id="CAC37765.1" /db_xref="GI:13923889" /translation="MDEGDRNGAEKKNFKLNKSKKKKKKKPKPYVSFMRYSN WLDKLYMVVGLAAITHGAGLPLMLVFGMDTDPFANAGNLGDLGALLENNNTSNIT DTVPVNLLEEDMTYAYVYSGIAGVLAAYIYVFWCLAAGRIHKKRQFPAHMR OEIGFVHDVGEINTRLTDVSKINIEIGKIGFPMASPTFGTFYGTGKMLT LVITLISPVGLSLAAVMAKILISPTDKRELLAYAKAGAAVEVLAIRTVIAPGOKKE LERNNKNEEAKRIGIKKATITANISGAFLIYASVLAFAFWTTLVLSKEYSTGOV LTFEVSILIGAFSVGQSPSTIEAFANAGAAFEIKIITDNKPSIDTSKSHKPDNIK GNLEFRNVEFSYPSRKEVYILKGLNLSQVQVLAIVGNSGKSTVOVMQRLYDPT EGWVSVDCMDIRITNVRFLREIIGVSSQEPVLAIVTAENTRREYEDVIMDEIKAVK EANVYDFMDLPOKEDTLVGERGALSGGQRIARIALRNKILLLDTSALDT ESEAVVOVALDKARKGRTTIVIAHRLSTRVADYIAGDDSVIYEKGHDELMKEGI YFKLTVMOTAGNEIELENADESEIDTETMSSHDSQSLRRSPRRSRSGCGD RKLSTKEALDSITPVSFWRMKILNTEMPYFVYGVPCATINGLQPPFAVETKIG LETRNDNAETKRONSNLFSLLFLVGLVSTFTFELGFTFGKAGEILTKRLRYVFS MLRQDVSWFDDPKNTGALTTTLRLANDAAVGAIGSLAITTONIANLGCIIISLY GMOETLLLLALPIITLAIAGVEMKLSGQALDKKELEAGKIAETAIENFRVYSLT OEKFEHMYDOSLQVYRNSLRKAHIEITFSPQAMMYFAGCFRGAVALVHSLM SFEDVLVFSNVFGMAVGOVSSFPADPYAAKAYSAHIIWIKETPLIDYSTEGIK PNTLEGNVTFNEVFNPTRLDITPVLOGLSLEVRKQOTLALVSGSGKSTVOVLLER FYDPLAGVLLDGEIKIQNLVOMLRAHIGISQEPILFDCSISNIAAGDSRVVSE EIVRAKANANHAFITSLIPNKYSTRVGCKTGLSGGQRIARIALRPOPHILLLDE ATSLADFESEKVEQALDKAREGRFCIYIAHRLSTIQVADLIIVFQNGRVHEGTHOO LLAKGIYFSNVSVQAGAKRQ"
BASE COUNT	1230 a 801 c 1039 g 1125 t
ORIGIN	
Query Match	34.4%; Score 710.8; DB 6; Length 4195;
Best Local Similarity	60.0%; Pred. 1.2e-140;
Matches 1261; Conservative	0; Mismatches 797; Indels 42; Gaps 3;
Oy	1 cgaagcaaggttcgactcaatcgtgtgtagcacaccgacttctactaactcgaatgacg 60
Dn	1835 CCAGAAAGAGGTGCGAACCCATTTGATAGCTCATCGTTTGTTCAGGTTCTAATGCGG 1894
Oy	61 atttgattgtgacccctaaagatggaatgctgtggcgagaagaagacatgctgaactaa 120
Dn	1895 ACGTATCGCTGCGTTTCGATGTGAGTCATGTGTGAGAAAGAAATCTGTGTAGCTCA 1954
Oy	121 tggcaaaagcaggtctatatactacttgt-----gatgcacagga 163
Dn	1955 TGAAGAGAGAAAGGCATTTCCTCAAACTTTGTCACATGACAGACAGCAAGTAATGAAATTG 2014
Oy	164 tattaanaagctgtatgaacagatgagatcaatgacatatctactgaagaagaagacaa 223
Dn	2015 AATTAGAAATGACAGCTGATGAATCCAAAGGAAATGATGATCTTGGAAATGCTTCA 2074
Oy	224 ctcaactctctgcacactctgtgaagaagatcaag-----taag 261
Dn	2075 ATGATTTCAGAGATCCAGCTAATTAAGAAAGATTCACCTCTAGAGATGCTCCGTGATCAC 2134
Oy	262 actcattgacaagcgtgaggaatccaccaatcctaagaagataagctctcctgaagctc 321



Db	1865	CCGAAAAAGGTCGGACACCCACATTGTGTATAGCTCATCTGTTTGTCTACAGTTCGTAATGCTG	1924
Qy	61	atttgatcttgagaccttaagaagatgaaatgctggcgagaaagagcaatgctgactga	120
Db	1925	ACGTCATCGCGTGGTTTGATGATGAGATCATTTGTGAGAAAGGAATCATGATGAAC	1984
Qy	121	tgagcaaaacgagtgctataltalttcacttgtagtgcacaggaatlaaaaaagctg	180
Db	1985	TGAAGAAGAAAGGCAATTTACTTCAAACTTGTCCACATGCAGACACAGGAATGAAGTTG	2044
Qy	181	aacagatgagatgaatgaactatctca-----ctgaagaagaaccaa	223
Db	2045	AATTAGAAAAATGCAGCTGATGAAATCCAAAAAGTAAATGATGCTTGGAAATGCTTCAA	2104
Qy	224	ctcaactctctcgaactctgtagagagatcaag-----tcag	261
Db	2105	ATGATTCGAAGTCCAGCTCTAATAGAAAAAGATCAACCTGTAGAGTGTCCCTGGATCAC	2164
Qy	262	aactcatgacaagcctgaaagaaaccaccaactaaagalaagctcttcctgaagct	321
Db	2165	AAGCCCAAGACAGAAAGCTTACGTACCAAAAGAGGCTCTGATGAAGATATACCTCCAGT	2224
Qy	322	ctcatataaaattttaaagttttaaacaagcctgaagagccttcttggtctggagaa	381
Db	2225	CCTTTGGAGAGATTAAGAACTTAATTAATTCGAATGAGGCTTAATTTGTGTGGTGTAT	2284
Qy	382	tgagctctgtctaaatlgagacgtctcatccagatattccatcactcttgcaaaatla	441
Db	2285	TTTGTGCAATTATTAATAGAGGGCTCGAACACACATTTGCATTAATATTTTCAAAATTA	2344
Qy	442	taaccatgcttbgaaataatgataaaaccaatlaag--cagatgcagaatlat	498
Db	2345	TAGGGGTTTTTACAAAGAAATGATGATCTGMAAACAAACAGACAGAAATAGTACTGTTTT	2404
Qy	499	ccatgatatctgcattatcttggtgtatttgccttgctgtagtatttcaagcaggaat	558
Db	2405	CACATATGTTTCTAGCCCTTGGAATTAATTTCTTTTATTCATTTTCTCTCGAGGGTTCA	2464
Qy	559	tttaagcgagcagcagggaaatttlaacgaatgaactgaagacacttgccctcaagaca	618
Db	2465	CATTTGGCAAAAGCTGGAGAAATCTCTCACCAAGGGCTCCGATACATGTTTTCCATCA	2524
Qy	619	tgatatcaggaatbtgcctgtgttgatlgaaagaaacagacaagcagaggtctgaca	678
Db	2525	TGCTCAGACAGAGATGTAGTTGGTTTGTATGATACCCCTAAAAACACACTGGAGCATTGACTA	2584
Qy	679	caatattagccatgatatatgacacaattcaagagcaacaggtccagaaatgcygctc	738
Db	2585	CCAGGCTCGCCAAATGATGCTGTCTCAAGTTTAAAGGGCTCTATAGTTCCAGAGCTTGTGTAA	2644
Qy	739	taacacaaatgcacaaactaactgaatgagacttcagttcaatcttccttataatlgatgg	798
Db	2645	TTTACCAGAAATATAGCAAACTTGGGACAGAGAAATATTTATTCCTTCACTATGTGTTGCC	2704
Qy	799	agatgacatctcgtatctgagatgattgctccagtaacttgccglacaggaatgactgaa	858
Db	2705	AACATAACACTGTACTCTCTGTAGCAATTTGTACCCATCATTTGCATATAGAGAGAGTTGTAA	2764
Qy	859	cccgacgaatgactggaatttggccaacaagataagcaagaacttaagcactgctgtaaga	918
Db	2765	TGAAAAATGTTGTCTGGGCAAGCACTGTAAAGATTAAGAAACACTAGAAAGTCTGTGGAGAA	2824
Qy	919	tacgaactggaagccttgagaaataacgtaactatgtagtcatlaacaaagggaaaagcct	978
Db	2825	TGCGTACTGAAGCAATAGAAAACCTTCGGAACGCTTCTTTCTTTGACTCAGGAGCGAAGT	2884
Qy	979	tcgagcaaatgtagaagagatgcttcacagactcaacacagaataactcgtagaagaagac	1038
Db	2885	TTGAAACATATGTATGCTCAGAGTTTGGAGGTACATACAGAAACCTTTGAGAGAAAGCAC	2944
Qy	1039	agattatggaagctgttatgtaactcaagcagtgcttataatttgcctatgcagcag	1098
Db	2945	ACATCTTTGGAAATTAATTTTCTCTCACCCAGGCAATGATGATTTTCTATCTGTGGAT	3004

QY	1059	ggtctcgatttggagcctatttaatlcaagctctgaaagaaatgaccccaagggcagtgctca	1158
Db	3005	gTTTTCCGGTTTGGAGCCCTACTTGGTGCCACATTAACCTATGACCTTTGGAGTGTTCGT	3064
QY	1159	taattttactgaattgcatactgagacatactggcaatcgaaacaacgccttgcttggtct	1218
Db	3065	TAGATTATTTTACGCTGTTGTCTTTTGTTGCCATGCGCCTGGGGCAAGTCATTATTTTGCTC	3124
QY	1219	ctgaatatctccaaagccaatatcg9ggctctg9gcattctgttgcctctgttctgnaaaagaaac	1278
Db	3125	CTGACTATGCGCAAGGCCAAAATATACACACGCCACATCATCATGATTCATTGAAAAAACCC	3184
QY	1279	caaatatagaacacgcgcgaatcgaagaagggaaaaagccagacacatctgnaaggaatttag	1338
Db	3185	CTTTGATTGACAGGTACACAGCAGGAAGCCCTTAATGCGCAACATTTGAAAGGAATAATGTCA	3244
QY	1339	agtttcggaagctctcttctctcatccatgctgcaccaatgttttcatccctccgtgct	1399
Db	3245	CATTTGTGGAGTTGTATTTCACATATCCACCGCAGCGGACATCCGAGTCTTCAGGGAC	3304
QY	1399	tatccctcagtatgagcgcgaggaagacagtagacattt9gg9agcagcgctctgga	1455
Db	3305	TGAGCCTTGAGGTGAAAGGAAGGGCCAGACGCTGGCTGTGGGGGCACACATGGCTGTGGGA	3366
QY	1459	aaagactctgcttcaactctctgaggaactttaabaaccccg9gaagaaacagctcgt	1518
Db	3365	AGAGACAGTGGTCCAGTCTCTGGAGGGTTTCTACGACCCCTTGGCGAGGAAGATGCTGC	3422
QY	1519	ttgagtg9tg9atgcaatgcaaaagaatttgat9tlaac9tg9ctc9gtlcccaaatagacaatcg	1578
Db	3425	TTGATGGCAAGAATATTAAGGCACTGAAATGTTCAATGAGTGGCTCGAGACACACTGGGGCATCG	3488
QY	1579	ttcctcaagcctctg9ctcttcaactctgcaactctgctgaaacatcgccctatg9taca	1638
Db	3485	TGTCCCAAGAGCCCATCTCCTGTTTACAGCAGCATTTCTGGAACATTTGCCCTATGGAGACA	3544
QY	1639	aaagcgcgtg9tg9ccattg9atg9atg9tcaaaagacgcgcaaatg9gaatatcatcat	1699
Db	3545	ACAGCCGGGTGTGTCTCACAGGAAGAGATTTGTGAGGGACGCAAGAGGCCAATACATACATG	3604
QY	1699	ctttattgaa9gtctccct9gaa9aat9acaa9aca9a9t9gaa9t9gaa9ga9gacagc	1758
Db	3605	CCTTCATCTGAGTCACTCGCTTAAT	3666
QY	1759	ttctctcgcgcgcagaaacaagaactagctatctgcaag9gctcttctccaaaaaccacaa	1818
Db	3665	TCTCTGGTGGCCCAAAACAAACAGCATTTGCCATPACTCTGTCCTTTTAAACAGCCCTCATTA	3724
QY	1819	ttttatgtt9gat9gaag9caacttca9ccct9gat9aat9aat9ca9t9ga9gt9gttctagc	1878
Db	3725	TTTTGCTTTTGGATGAAAGCCACGTACACTCTGGATACAGAAATGAAAGGTTGTCCAAAG	3788
QY	1879	atgcctctgataaa9ccag9ac9g9ga9ga9cat9t9c9tag9tg9tca9tca9cag9ctctcg	1938
Db	3785	AAGCCCTTGACAAAGCCACAGAAAGAGCCGACACTGTGACATTTGGATTGCTCACCCCGTGTCCA	3844
QY	1939	caattcgaacgcgcgaattg9atg9tctctgcaaatg9ga9aa9ga9ga9ga9ga9aa	1999
Db	3845	CCATCCAGAAAGTCAGACTTATATATGTTGTTTTCAGATGCGACAGTCAAGGAGCATTTGCA	3904
QY	1999	cttcaatcaag9cctctct9ga9aat9ga9acata9tat9tla9t9tag9t9ga9tg9ac9tagc	2055
Db	3905	CGCATACGACACTGCTGGCACAGAAAGCATTAATTTTTCATATGTCATGTCACAGCTGC	3966
RESULT 15			
AR051650			
LOCUS	AR051650	4264 bp	DNA
DEFINITION	Sequence 5 from patent US 5830697.		PAT
ACCESSION	AR051650		
VERSION	AR051650.1	GI:5975014	
KEYWORDS			
29-SEP-1999			

	RESULT	15			
ARO5I650	LOCUS	ARO5I650	4264 bp	DNA	PAT
DEFINITION	Sequence	5 from patent US 5830697.			29-SEP-1999
ACCESSION	ARO5I650				
VERSION	ARO5I650.1	GI:5975014			
KEYWORDS	.				



Search completed: April 22, 2002, 22:44:24  
Job time: 27172 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:44:24 ; Search time 14771.4 Seconds

(without alignments)  
3189.671 Million cell updates/sec

Title: US-09-873-409-10

Perfect score: 2856  
Sequence: 1 cctaatctctcaataatctc.....atgcacagtcagtcagtcaga 2856

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_ov:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rnd:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rnd:\*  
36: em\_htg\_other:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1030.2	36.1	3988	6 AX024454	AX024454 Sequence
2	1030.2	36.1	4192	9 AF016535	AF016535 Homo sapi
3	1030.2	36.1	4646	6 I49610	I49610 Sequence 2
4	1030.2	36.1	4646	9 HUMMDR1	M14757 Homo sapien
5	1030.2	36.1	4659	6 I08557	I08557 Sequence 3
6	1030.2	36.1	6505	6 AR028671	AR028671 Sequence
7	1030.2	36.1	9318	6 AR028672	AR028672 Sequence
8	1028.6	36.0	4264	6 AR051647	AR051647 Sequence
9	1028.6	36.0	4264	6 AR051650	AR051650 Sequence
10	1028.6	36.0	4378	6 E02326	E02326 Multidrug r
11	1028.6	36.0	8630	6 AX012321	AX012321 Sequence
12	1027.2	36.0	3924	6 AX024455	AX024455 Sequence
13	1027.2	36.0	3924	9 HUMMDR3	M23234 Human membr
14	1027.2	36.0	8630	6 AX012320	AX012320 Sequence
15	1026.6	35.9	4279	6 AX105057	AX105057 Sequence
16	1026.6	35.9	4279	6 AX105078	AX105078 Sequence
17	1026.6	35.9	4279	6 AX105080	AX105080 Sequence
18	1025.4	35.9	4669	6 AR055785	AR055785 Sequence
19	1025	35.9	4045	12 AF269224	AF269224 Synthetic
20	1025	35.9	4279	6 AX105082	AX105082 Sequence
21	1023.4	35.8	4317	4 AF045016	AF045016 Canis fam
22	1023.4	35.8	4317	6 AX105059	AX105059 Sequence
23	1022.2	35.8	4669	6 AR091275	AR091275 Sequence
24	1022	35.8	3987	10 CRUPGP1185	M59254 Chinese ham
25	1022	35.8	4296	10 CRUPGP1	M60040 C.griseus p
26	1022	35.8	4304	10 CRUPGP1165	M59253 Chinese ham
27	1021.6	35.8	4018	5 GGA9729	AA009799 Gallus ga
28	1020.6	35.7	4186	6 AX108654	AX108654 Sequence
29	1020.6	35.7	4195	6 AX108656	AX108656 Sequence
30	1009	35.3	4280	10 CRUPGP11	M60041 C.griseus p
31	981.8	34.4	4084	10 MUSMDRA	I03398 Mouse mdr g
32	979	34.3	3858	4 OU078609	U78609 Ovis aries
33	969.4	33.9	4927	10 AF257746	AF257746 Rattus no
34	969	33.9	3905	10 CRUPGP111	M60042 C.griseus p
35	966.8	33.9	4323	10 AF286167	AF286167 Rattus no
36	953.2	33.4	4233	6 AR123273	AR123273 Sequence
37	950.2	33.3	3682	10 MUSPGLY1A	M24417 Mouse phosp
38	950.2	33.3	4924	10 MUSMDR1A	M33581 Mouse p-gly
39	949.6	33.2	4298	10 MUSMDR	M14757 Mouse multi
40	948.6	33.2	4356	10 MUSMDRAA	M30697 Mouse multi
41	936.8	32.8	3912	10 RATPGLYCO	L15079 Rat p-glyco
42	932.8	32.7	3489	4 AB029153	AB029153 Felis cat
43	926	32.4	4254	10 RATMDRM	M81855 Rat mdr mRN
44	925.4	32.4	4390	5 XLU17608	U17608 Xenopus lae
45	856	30.0	2726	6 I33621	I33621 Sequence 1

#### ALIGNMENTS

RESULT 1  
AX024454  
LOCUS AX024454 3988 bp DNA  
DEFINITION Sequence 1 from Patent DE19857750.  
ACCESSION AX024454  
VERSION AX024454.1 GI:10184622  
KEYWORDS  
SOURCE  
human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 3988)  
AUTHORS Stresemel, W.  
JOURNAL Patent: DE 19857750-A 1 24-FEB-2000;  
STRESEMEL WO/FGANG (DE)  
FEATURES  
source  
1..3988  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="MDR1/menschliche MDR-Genfamilie"





Db	3208	CCGGACATCCAGCTGCTTCAGGAGCTGAGCCCTGGAGGTGTAAGAAGGCCACAGCGTGGCT	3267
Oy	2224	tttctggggagcagcgcgtctgtggaaagcacctctcgtcaactctgcagagacttat	2283
Db	3268	CTGGTGGGCGACGACGCGCTGTGGGAAGACACAGTGGTCCACCTCTGGAGCGGTTCTAC	3327
Oy	2284	gaccccgycgaagagacaagtgcctgttatatgtgtgagatgacaaagaatgtgaatctaacg	2343
Db	3328	GACCCCTTGGCAGGGAAGTGTCTGCTTGTATGGCAAGAATAAAGCAGCTGAATGTTCCAG	3387
Oy	2344	tgtctcgcttcccaatatgcaalcgctccctcaagaagcctgtgctctcactgcagcagcalt	2403
Db	3368	TGGCTCCAGACGACACCTGGCGATCCTGTGCCAGGAGCCCATCCTGTTGACTGCAGCATT	3447
Oy	2404	gctggagaacatcgctctatgtgtacaaagccgtgtgtgtgacatagatgagatcaagaag	2463
Db	3448	GCTGGAACATTGCTTATGAGACACACAGCCGGGTGCTGTCCAGAGAGAGATTGTGAGC	3507
Oy	2464	gccgcaaaatgcagcaaatatccatctctttatgtgaagctccctcgtgagaataacaaca	2523
Db	3508	GCAGCAAGAGAGGCCCAACATACATGCCTTCATCGAGTCACTGCCATAATTAATATAGCACT	3567
Oy	2524	caagtgtgacgtgaaaggagcacagcttctctgtgcgcgaagaaacaaagactagctattgca	2583
Db	3568	AAAGTAGGAGACAAAGGAACTCAGCTCTCTGTGGCCAGAAAACAACGCATTCCTAGCT	3627
Oy	2584	aggagctcttcccaaaaaccccaaatlttatgtgtgtgtgagtgagccacttcaacccctgat	2643
Db	3628	CGTGCCCTGTTTAGACACAGCTCATATTTTGGCTTTTGGATGGAAGCCACGCTCTTGAT	3687
Oy	2644	aatgacagctgaaaggtgtgttcaagcatgcctctgataaagccagagacggaagaagatgc	2703
Db	3688	ACAGAACTGTGAAAAGGTTGTCTCCAAAGACCCCTGGCAAGACAGAGAGGCCGCGACCTGC	3747
Oy	2704	ctagtggtcactcacaaaggctctctgcaatttgaaagcgaagatttgatagtggtctgcag	2763
Db	3748	ATTGGAATTGCTCACCGCCTGTCCACCATCCGAATGCAAGACTTAATGTGGTGTTCAG	3807
Oy	2754	aatggaagaatataaagaaacaaagaaactatcaagaagctctctgagaatgcataatata	2823
Db	3808	AATGCACAGTCCAAAGACATGCGCACGATCAGCAGCTGCTGGCACAGAAAGCATCTAT	3867
Oy	2824	tttaagtagtgaatgacagatgcag	2848
Db	3868	TTTCAATGTCAGTGTCCAGGCTG	3892
RESULT	2		
AF016535	LOCUS	AF016535 4192 bp mRNA PRI 03-SEP-1997	
DEFINITION	Homo sapiens P-glycoprotein (mdr1) mRNA, complete cds.		
ACCESSION	AF016535		
VERSION	AF016535.1 GI:2353263		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 4192) Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B.		
AUTHORS	Internal duplication and homology with bacterial transport proteins in the mdr1 (P-glycoprotein) gene from multidrug-resistant human cells		
TITLE	CELL 47 (3), 381-389 (1986)		
JOURNAL	CELL 47 (3), 381-389 (1986)		
MEDLINE	87028230		
REFERENCE	2 (bases 1 to 4192) Chen,C., Duran,G.E., Steger,K.A., Lacayo,N.J., Jaffrezou,J.P., Dumontet,C. and Sikic,B.I.		
AUTHORS	Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein, altered phenotype, and resistance to cyclosporins		
TITLE	J. Biol. Chem. 272 (9), 5974-5982 (1997)		
JOURNAL	J. Biol. Chem. 272 (9), 5974-5982 (1997)		
MEDLINE	97190336		

REFERENCE	3 (bases 1 to 4192)
AUTHORS	Chen, G., Lacey, N. J., Steger, K. A. and Sikic, B. I.
TITLE	Direct Submission
JOURNAL	Submitted (28-JUL-1997) Medicine, Stanford University School of Medicine, Stanford, CA 94306, USA
FEATURES	Location/Qualifiers
SOURCE	1..4192
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7q21.1"
	/cell_type="drug-resistant uterine sarcoma"
	/cell_line="MES-SA/DXp"
	/clone="8.1"
gene	1..4192
	/gene="mdr1"
	/note="multidrug-resistance gene"
	120..3959
CDS	/gene="mdr1"
	/function="transporter protein"
	/note="multidrug-resistance"
	/codon_start=1
	/product="P-glycoprotein"
	/protein_id="P335326"
	/db_xref="GI:2353264"
	/translation="MDLEGDRNGAKKKNEFKLNKSEKDKERKPTVSVFSMRYSNWLDLTVAVGTLLAIINGAGLPLMLVFGEMTDTFFANAGNLEDLMSNITRSDINDNGIFPNLEEDMPTRYAYYSIGAGVLAAYIGVSPGCLAGAQRIKIRKROFPAIMROBIFGMFVHVHVGELNTRLTDDVSKINGIDDKIGMFQSMATPEFTGIVGTFGWLKTLIVT LAISPIVGLSAAVAKILSTSTDELLAYAKAGVAEYVLAIFTVIARFGCKKELERYNKLEAKRIGIKKATITANISIGAAFLLIYASTALFWGTGLVLSGEYSICQVLVFSVLIGAFVWGQASPSIEAFNARGAAVEYFKIIDNKPISIDYSKSGHKDNIKGNLEFRNDFYSPSKREKVIKGLKLVQSOQVALVNSCGSKSTVOLMQRLYDTEGVNSQDQDIRTINVRFLREIIGVSOEPLFATLIENTRGRNENTMDEIKAKEANAAYDFIMKLPHKEDTLVGERGALSGGQKORIAIARALRNPKLILDEATSLPESPAVVOALDKARKGRTIYIAHRLSTVRRNADVIAGDDVYIEKGNHDEMEKGIYFKLVOTQAGNEVELENADESKSEIDALESSMDSRSSLIKRSTRSVRSGQADQRKISTREKLADSIPEVSEFMFLMKLNLTEMPYFVGVFCALINGLQAPAFALIFKRIIGVFLRIDDEPKRONSNFSLFLALGIIISFTFFLOGFTFGKAGILTKRLRYVWFMSLRODSWDEPKNTGTGALTTRLANDAQQVGAIGSRVLATVONIANGTGIIISFTYGMQLTLLIALLPIIAGVVEMKLSGQALDKKKELEGACKIATEAIENTRVSLSIOEOKPEHHYAGSLQVPRNSLRKAIHIGITSEFQAMMYPSYACGFRGATLVANHKLMSPEDVLVFSAVFGAMAVGVSSFPADYPAKAKISAHTIMITEPTPLIDSTIEGKLPNLLEGNVTEGEVFNYPTRPDIPVLOGLSLEVKKGQTLAVSSGCGKSTVOLLERFPDL LAGVLLIDGKEIKRLNQWMLRAHIGIYSOEILFDCSIAENIAGDMSRYVSOEIVRLAKRANIHAFIESLPNKYSTVKGKQSGQLSGGQKORIAIARALROPHILILDEATSLADPEKRYVEALDKARKGRCIYIAHRLSTVRRNADVIIVFQNGRVAKHGTHQQLLAQKGIYFSNVSQAGTKRQ"
variation	659
	/gene="mdr1"
	/citation=[1]
	/replace="c"
	673..674
variation	/gene="mdr1"
	/citation=[1]
	/replace="t"
	1351
variation	/gene="mdr1"
	/citation=[1]
	/replace="c"
	1428-1429
variation	/gene="mdr1"
	/citation=[1]
	/replace="ttc"
	3551
variation	/gene="mdr1"
	/citation=[1]
	/replace="c"
BASE COUNT	1243 a 1028 g 1124 t
ORIGIN	797 c



[illegible][illegible]

Oy	1324	gtacgtattctaacgaaggaattattttccgcgcagagcagaggcggaattttacgaataga	1380
Db	2726	ATTTCATATTTTTCCTCAGAGGTTTCACATTTGGCCAAAGCTGGAGAACTCTCACCAAGCG	2785
Oy	1384	ttaaagaacttgcgccttcaagaagccatgattatcacgaatatthccctggttttcagtlaaag	1443
Db	2786	CTCCGATACATGGTTTTCCATTCATTCGCTCACAGAGATGTGATTTGGTTTGAATGACCT	2845
Oy	1444	gaaaacagcagcagggagctgtgacaaatatatagccatagatatatgcacaaattcaagga	1503
Db	2846	AAAAAACCCACTGGAGCATTTGACTTACACAGCTCCGCCAATGATGCTGCTCAAGTTAAAGG	2905
Oy	1504	gcaacaggttccaggaattgcgtcttaacacaaatgcacaaatgcacaaatgcgaacttcaagt	1563
Db	2906	GCTTATAGTTCACAGCCTGCTGCTGATTTACCAGCAATATATAGCAAACTTTGGACAGATA	2965
Oy	1564	atcattcccttatalatatgatacggagcagatccctgattcagatattgcttccagta	1623
Db	2966	ATTATATCTTTCATCTATGTTGGTGGCAACTACATGTTACTCTTATGCAATGTATGCCATC	3025
Oy	1624	cttcgcgtgcaggaatgatattgaaccgcgaacgaatgactgagtttgcacaacaaatagaag	1683
Db	3026	ATTTCATATAGCAGAGAGTTGTTGAAAGAAATGTTGCTCGCAAGCAACTGAAGATATAG	3085
Oy	1684	caagaactlaagcactgcctgtaaaagatagaacatgaaagcttttgagaaatatacgtactata	1743
Db	3086	AAAGACTAGAAAGTCTCGGAGATTCGCTACTGAAAGCAATATGAAACATTTCCGACCGTT	3145
Oy	1744	gtgtcatttaacaagcagaaagaccccttcgcgcgaacgatgatagaagaatgcttcagactcaa	1803
Db	3146	GTTTCTTTTGATCTAGGAGCGAAGATTGTGAACATATGTATGCTACAGATTTTGCAGGTACCA	3205
Oy	1804	cacagaataactcctgcagaagaacacagatatattggaagctgttaacatcaacatgccc	1863
Db	3206	TACAGAAACCTTTGAGAGAAAGCACACATCTTTGGAATTTACATTTTCCCTCACCCAGCA	3265
Oy	1864	tttatataatttgcctatgcacgaagggttcgaatttggagcctatttaatcaagctgga	1923
Db	3266	ATGATGATATTTTTCCTATGCTGTGATGTTTCCGGTTTGGAGCCTACTTGGTGACATATAA	3325
Oy	1924	cgaatgaaccccaagggcagctgttcaatagtttttaacgcaatgcgaatagagacatacgtgc	1983
Db	3326	CTCATGACCTTTGAGGATGTTCTGTATAGATTTTTCAGCTGTGTCTTGTGTGCCATGGCC	3385
Oy	1984	atcggaaaaaacgcctcglttttgcctccgtaabattccaaagccaaatcggggctgcgat	2043
Db	3386	GTGGGGCAAGTCATGTTATTTGCTCTCGATATGTCAAAGCCAAATATACACACCCAC	3445
Oy	2044	ctgtttgcctgttggaaaagaacaaacaaatatagaagcgcgaagcagaagaagggaagaag	2103
Db	3446	ATCATCATGATCATTTGAAAAAACCCCTTGATTTGACAGCTTACAGCACGAGAAAGCCTATAG	3505
Oy	2104	ccagacacatgtaaggaagaatttagagtttcggaagatctcttctcattccatcagtcgc	2163
Db	3506	CCGAACACATTTGAAAGGAATATGTACATTTTGGTGAAGTTGTATTCATATCCACCCGA	3565
Oy	2164	ccagatgttttcaatccctccgtgtgcattatccctcaagtattgagcgcaggaagaagacagtagca	2223
Db	3566	CCGGACATCCCAAGCTTTCAGGGACATGAGCTTGAGAGTACAAAGAGGCCAGCGCTGGCT	3625
Oy	2224	ttttgtgggagcagcggcctgtggaaaagaacttgcctcaacttctgcagaagcattat	2283
Db	3626	CTGTGGGGCAGCAATGGCTGTGGGAAGAGCACAGTGTGCTCAGCTCTCGAGGGGCTTCTAC	3685
Oy	2284	gaacccctgcagaagcaagatgctgttattgatggtgtgagatgcacaaagaattgaatgtagc	2343
Db	3686	GACCCCTTGGCAGGAAAGTGTGCTTGATGATGCACAAAGAAATTAAGAGCACTGAATGTTCAG	3745
Oy	2344	tgtcgtcgtttcccaaatagaatcagatcgttcccaagagcctgctcttcaacagcagcatt	2403
Db	3746	TGGCTCCAGACACACTGTGGCATCTGTCTCCAGAGACCCATCCTGTTTATGCTGCAGCAATT	3805

OY	2404	gctbagaacbtccctatagtttgacaaacagccgttgcgttgcattagatgagatcaaa	2463
Db	3806	gctbagaacattgctctatgtagacaaacacagccgggtggtctcacagaagagatcgt	3865
OY	2464	gcgcgaatbgaacaaatalccattcttattbgaagctccctgagaaatacaaca	2523
Db	3866	gcacccaagggagcccaacatcacatgccttcacacgaatcactcaataaataatgac	3925
OY	2524	caattbgaactgaaagagacacagcttcttcgcgcgagaaacaagaagctgcatgca	2583
Db	3926	aaattagagagacaaaggaaatctacgctctcgtggggccaaataaacgattgcc	3985
OY	2584	agggcctctcccaaaacccaaatlttatbgttgatgagagccacttcagccctgat	2643
Db	3986	cgctccctctgtttacacgcctcatatttttgcttttgatgaagccacgctacgct	4045
OY	2644	aatgacagtgagaagtgtgttcgacatgcctctgataaaagccagagcgggaagaacat	2703
Db	4046	acacaaagtgaaaaggtttgttccaaagagccctctggacaaagccagaaagccgc	4105
OY	2704	ctatgtgtcaactcaagcctctctgcgaattccagaacagagatttatagttgtc	2763
Db	4106	attgtgattgcttcacccctctctccaccatccgaatgcagacttaattagttgt	4165
OY	2764	aatggaagaataaagagacaaggaactcatcaagagctccctgagaatctgagacata	2823
Db	4166	aatggcacagtcgaagagacatgtgcacgacatcacacactcctggcacagaaagc	4225
OY	2824	tttaagttaagtgaatgcaagctag	2884
Db	4226	ttttcaatggtcagttgcacagcctg	4250

	RESULT	4
HUMMDR1 LOCUS	4646 bp mRNA PRI	03-DEC-1999
DEFINITION Homo sapiens P-glycoprotein (PGYI) mRNA, complete cds.		
ACCESSION M14758		
VERSION M14758.1 GI:187468		
KEYWORDS P-glycoprotein; drug resistance protein; transport protein. Human drug-resistant cell line KB-C2.5 cDNA to mRNA, clones Lambda-HDR1[O, 5, 104].  		
SOURCE Organism Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE 1 (bases 1 to 4646) Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B. Internal duplication and homology with bacterial transport proteins in the mdrl (P-glycoprotein) gene from multidrug-resistant human cells		
TITLE JOURNAL Cell 47 (3), 381-389 (1986)		
MEDLINE REFERENCE 87028230		
AUTHORS Ueda,K., Clark,D.P., Chen,C.J., Roninson,I.B., Gottesman,M.M. and Pastan,I.		
TITLE The human multidrug resistance (mdr1) gene. CDNA cloning and transcription initiation		
JOURNAL J. Biol. Chem. 262 (2), 505-508 (1987)		
MEDLINE REFERENCE 87109132		
AUTHORS 3 (bases 971 to 985; 3095 to 3109) Kioka,N., Tsubota,J., Kakehi,Y., Komano,T., Gottesman,M.M., Pastan,I., and Ueda,K.		
TITLE P-glycoprotein gene (MDR1) cDNA from human adrenal: normal P-glycoprotein carries Gly185 with an altered pattern of multdrug resistance		
JOURNAL Biochem. Biophys. Res. Commun. 162 (1), 224-231 (1989)		
MEDLINE REFERENCE 89322246		
AUTHORS [2] sites. Draft entry and computer-readable sequence [1] kindly submitted by T.B.Roninson, 13-AUG-1987.		
TITLE The sequence shown is of a cDNA clone initiating at a minor upstream transcription initiation site and containing the major		
JOURNAL COMMENT		

FEATURES                    site of transcription initiation.  
Source                    Location/Qualifiers  
                            1. 4646  
                            /organism="Homo sapiens"  
                            /db\_xref="taxon:9606"  
                            /map="7q21"  
                            1. 4646  
                            /gene="PGY1"  
                            382  
                            /gene="PGY1"  
                            /note="g in [1]; a in [2]"  
                            425. 4267  
                            /gene="PGY1"  
                            /note="P-glycoprotein"  
                            /codon\_start=1  
                            /protein\_id="AA59575.1"  
                            /db\_xref="gi:307180"  
                            /translation="MDLEGDRNGGAKKNEFKLNKSEKKKEKPPYVSEPMFYSN  
MDKLYVVGFLAIIHGAGLPLMIVGEMTDIIPANGNLEDLSNTNDSINDTG  
PKNLEPDMPTAVYSGIGAGVVAAYIQVSPICAGAGRIHKIRKQFPAIMROEI  
GMEDVHDVGEINLRLDVSKEINVEYIGDKIGFQSMATFEFTGTVGTRKMLTLVI  
LAISPVGLSAVNAKILISFTDKELLAYAGAVAEEVLAIRTVIAFGQKELER  
YKNMLEARKRIGKAITANISIGAFLIYASVALFMYGTTVLSEGVIGVLT  
FFSVLIGAFSVGQASPIEAPANRGAAYELFKIIDNPSIDSYKSGHRDNKGNL  
EFNVHFSYPSRKEVKILKGLNKVSGQTVAGLNSGCGSTVYOLMRLYDPEGM  
VSVDGDIRTNVRFLEITIGVSOEPLVFTTAENIRYRENVMTDELKAYKEAN  
AYDTMKLPKFDTLVGBRQALSGGQKRIARALRNKRLILDEPAISALDTESE  
AVQVALDKARKGRITVIAHRLSTVRNADVIAGDDVIEKGNHBLMEKGYFK  
LVTMOTAGNEVELNADSEKIDALEMSNDSRSLIRKSTRSRVSGQADRLK  
STKALDESIPIVSEFWRIMKLNTEMPEYVGVCAIINGLOPAFALIFSKIGVFT  
RIDDEPRKRONSLFLLFALGIISFTFPLQGTGKAGEILLKRLRYVPMRLR  
ODVSMDDPKMTGALTFLRLANDAOKVAGISRLAVTTONIANGCIIISFYQO  
LTLILAIPITAIAGVYEMKLSGALKKKKEGACKIATEINERTIVSLTQEO  
KEFMYAOSLOVPRNSLRKAHIFGTFSTQAMMYFACCFRGALVAHKLMSEF  
DVLVFSAVYAGAMVQVSSAPDYAKIASAHIMIEKTPILDSYSTEGLEPNT  
LEGWTFGEVFNPTREDPIVLOGISLEVKGOTLALVSSGCGSTVYOLLEPFD  
PLAGKVLDEKIRLNYOMLRHGLIYSOBPLFEDCSIAINIAVGDNSRVSOEIV  
RAAKEANIHAFESLPNKYSRKVGDKQOLSGGQKRIARALRYROPHIILLDEATS  
ALDRESEKVOEALDKAREGRTCIYIAHRLSTIONADLIIVFQNGRYVEHGHOOOLA  
OKGIYFSKVSVOAGTKRO"  
                            964  
                            /gene="PGY1"  
                            /note="t in [1]; c in [2]"  
                            978. 979  
                            /gene="PGY1"  
                            /note="t in [1]; ga in [2] Val->Gly"  
                            1660  
                            /gene="PGY1"  
                            /note="c in [1]; t in [2]"  
                            2065  
                            /gene="PGY1"  
                            /note="c in [1]; t in [2]"  
                            3101  
                            /gene="PGY1"  
                            /note="g in [1]; t in [2] Ala->Ser"  
                            3859  
                            /gene="PGY1"  
                            /note="c in [1]; t in [2]"  
                            4460  
                            /gene="PGY1"  
                            /note="a in [1]; g in [2]"  
                            /note="a in [1]; g in [2]"  
BASE COUNT    1371 a    892 c    1129 g    1254 t  
ORIGIN        154 bp upstream of Avar site; chromosome 7q21.1.

Query Match                    36.1%; Score 1030.2; DB 9; Length 4646;  
Best Local Similarity    62.2%; Pred. No. 3e-198;  
Matches 1695; Conservative 0; Mismatches 988; Indels 42; Gaps 3;

Qy    166 attattgataaagaccagatataaactttccacagctgagatataaactgatacc 225  
          ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1526 ATAAATTGATAAATAGCAGATATTGACAGCTATTGCAAGATGGGACAAACAGATTAAT 1585

Qy    226 atagagaagactgtagaattaaatgcttctcaattacacatcaaacatctatc 285  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1586 ATTAAGGGAAATTTGGAAATTCAGAAATGTCACCTCAGTTACCCATTCGAAAAGAAAGTT 1645

Qy    286 aagattcgaagagctcgaatccgaatlaagtctgagagacagctgccttggctg 345  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1646 AAGATCTGGAAGGCGCTGAACTGAAAGTGCAGAGTGGGCAACAGTGGCCCTGGTTGGA 1705

Qy    346 ctcaatgagcagtgagaaagatagcagtgatccagcttctgcaagagttatagatccgaat 405  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1706 AACAGTGGCTGGGGAAGAGCAACACAGTCCAGCTGATGCAAGAGCTCTATGACCCCA 1765

Qy    406 gatggcttatcatgtagatagaatagacatcagagctttaaagtgtgcatctatga 465  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1766 GAGGGAGTGGTCAGTGTTGATGACAGAGATATTAGACCATTAAGTTAAGTTTCTACGG 1825

Qy    466 gaccatctgagatgtagtaagaagccgttcttctgaggaaccacatcagtaaacat 525  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1826 GAAATCATTGCTGTGTGTGATGACAGAACCTGTATTGTTGCCACACGATAGCTGAAGC 1885

Qy    526 atcaagatgagcagagatgtagtgcctgaatagaagaatgagagagcagcaaggagca 585  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1886 ATTCCTATGGCCCTGCAAAATGTCAACATGATGAGATTGAAGAAAGCTGTCAAGAAAGCC 1945

Qy    586 aatgagtatgatttataatcagtgagcttccataaattatcatattgtagggaaaaa 645  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    1946 AATGCTATGACTTATTATCATGTAAACTGCTCTATTAATTTGACACCTGTGTAAGAGACA 2005

Qy    646 ggaagctcaaatgagtgagagggcagaagaacagagatcgcgaatgctgcttgaatcga 705  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2006 GGGGCCAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2065

Qy    706 aaccccaagatctgaattttagatgagctacgtctgcctgagatcgaagaagatga 765  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2066 AACCCCAAGATCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2125

Qy    766 gctgttcaagctgcagctgagagagcagcaagagctgcagatcgaatcgtgtagaacac 825  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2126 GTGCTTCAAGTGGCTCTGTGATTAAGCCAGAAAGGTGGACCAACCTTGTGATAGCTCAT 2185

Qy    826 cgacttctactatctcgaagtgcaagatttgatgtgacccctaaagatggaatgctgagc 885  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2186 CGTTGTGTACAGTTCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2245

Qy    886 gagaagaagagcagatgctgaactaatgagcaaaacgaaggtctatattatctgtagatg 945  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2246 GAGAAAGGAATATCATGATCACTGATGAAGAGAAAGGCACTTACTTCAAACTGTGCACA 2305

Qy    946 tcacagatattaaanaaagctgataacagatgagagatgacatattcta----- 997  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2306 ATGCAGACAGCAGGAATGAAGTTGATTAAGAAATGCGAGCTGATGATGATGATGATGATG 2365

Qy    998 -----ctgaagaagaagacacatcctcctcctcctcctcctcctcctcctcctcctcctc 1045  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2366 ATTGATGCTTGGAAATGCTTCAATATGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 2425

Qy    1046 -----agtcagactcattgacaagaagctgaggaatccacccaatct 1086  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2426 ACTCGTAGAGAGTCCGTGATGATCAGAACCCAGAGAAAGCTTATGATCAAGAGAGCT 2485

Qy    1087 aagaagataagctctcctcgaagctcctcctcctcctcctcctcctcctcctcctcctc 1146  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2486 CTGGATGAAAGATACCTCAGCTTCTCTTGGAGGATTAAGCTTAATTAATTAATTAATTA 2545

Qy    1147 tggccttctgctctcctgagacatgctcctcctcctcctcctcctcctcctcctcctcctc 1206  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2546 TGGCCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2605

Qy    1207 ttctccatcctcttgcaaaaatataacacatgcttggaacaataatgataaaac-----aca 1263  
          ||| ||||| ||||| ||| ||||| ||| ||||| ||| ||||| |||  
Db    2606 TTTGCAATATATATTTTCAAAAGATTAATAGGGGTTTTCACAAAGATGATGATGATGATG 2665

Qy    1264 ttaagcatgatgacgaagaatttatccatgatatatcgtcaatttgggtgttatcttggctt 1323





OY	346	ctcaagtcgaagtcggaagaagtacgtatgcacgacttctcgaagttatatgatccgat	405
Db	1706	AACAGTGGCTGTGGGAAGACCACAACAGTCACACTGTAATCAAGAGCCTCATATCACCCCA	1765
OY	406	gatggcttacatcaygygagtcgaagatgcataccaagctttaaattgtyggcatatcga	465
Db	1766	GAGGGGATGGGCACAGTGTTGATGGACAGAGATATTAGACCATATAATGTAAAGTTTCTACGG	1823
OY	466	gaaccatctggagtgyltiagtcgaagacctgtlittgltcggaaccaacaatcaatcaat	525
Db	1826	GAATTCATTTGGTGTGGTAGCTACAGGAACCGTATTTGTTGCCAACACGATACCTGAAMAC	1885
OY	526	atcaagatatygaagatgtagtctgtactgtatgaaagataatgaaagagcaagaaagca	585
Db	1886	ATTGCGTATGGCCGTGMAAATGTCCACCATGGATGAGATTTGAAMAAGCTGTCAAGAAGCC	1945
OY	586	aatgcgatatttatcacatcygagttcccataaatatatacatattglttagggaaa	645
Db	1946	AATCCCTATGACTTTTATCATGAAACTGCTCATAAATTTGACACCCCTGGTTGGAAGAGA	2005
OY	646	ggaagctcaaatgagtcyagygccagaacacagagatccgaattgctgcgttcccttaatcga	705
Db	2006	GGGGCCAGTTGAATGGTGGGACGAAAGCAGAGSANTGCCATTGCACGTGCCCTGGTTGCG	2065
OY	706	aacccaagaatcttgatatttagatgtagcgaogtctgcgccgtatccaagaacagica	765
Db	2066	AACCCCAAGATCCCTCCGTCGTGAGTGAAGCCACGTCAGCTTGAGCAACGAAGAAGCA	2125
OY	766	gctcttaagctgcacgcggsgaaggcgagcaaaugtcggatcaacaatcgttgtagcaac	825
Db	2126	GTGGTTCAAGTGCTCTGGATTAAGGCCAGAAAAGTGCGGACCAACCATTTGTGATACCTAT	2185
OY	826	cgaacttctaactlcgaagtctgaattgagattgtatgtaccctaaagatgaaatgctgcg	885
Db	2186	CGTTGTCTACAGTGTGATATGCTGAAGTCATCGCTGCTTGATGATGATGATCTATTGTG	2245
OY	886	gagaaagagcaacatgctgaacataatgcaaaaaagagctatatlactactgtgatg	945
Db	2246	GAGAAAGGAATCATGTATGTAACCTCAITGAAGAAAGAAAGGATTTACTTCAAATCTGTACA	2305
OY	946	tcaacagatatlaaaaaagctgtatgaaacagatgtagtcaattgacatalctca-----	997
Db	2306	ATGCGACACAGCAGGAAATGGAATTTGAATAATGCAAGCTGATGAATCCAAAGTGA	2365
OY	998	-----ctigaagaagaacccaactcaactctccctgcacactctgaaagacatca---	1045
Db	2366	ATTGATGCCCTGGAAATGCTTCAATTAATGATCAAGATCCAGTCTATTAAGAAAAAGATCA	2425
OY	1046	-----agtcagaacttaatgacaagagctgaaatcccaacatct	1086
Db	2426	ACTGCTAGAGATGTCCTGTGATACAAAGCCAAAGCAGAAACCTTAGTACCAAAGAGCT	2485
OY	1087	aagaagatagctcttcctgaagctctctcataaaaaatttaagttaaacaagccgaa	1146
Db	2486	CTGAGTAAAGTATNACTCCAGATTTCTTTTGGAGATTATTAACCTAAATTTAATCTAA	2545
OY	1147	tggagcttgltgltccggygggaactggcgtctgtctciaaaggaagacgttcatcagta	1206
Db	2546	TGGCCTTATTTTGTGTGTGTATTTTGTGCACTTATTAATGAGAGCCTGCACACACGA	2605
OY	1207	ttttccatcaacttgcgaanaatataaacatglttgcgaataatgataaaaac---aca	1263
Db	2606	TTTTGCATATATTTTTCAAAGATTATAGGGGTTTTTACAAAGATTGATGATCCTGAACA	2665
OY	1264	ttaaagatgatgcagaanaattatccacagatatccgtcaatttgggtgtatattgctt	1323
Db	2666	AAACGACAGATAGTAACTTTTTCATATTTGTTCTAGCCCTTGGAAATTTATTTCTTTT	2725
OY	1324	gtcagctattcatcaycgaggtattattcgaagcgagagcgaggaattttaacgaatga	1383
Db	2726	ATTATCATTTTTCCTTCAGSGTTTCACATTTGGCAAAAGCTGGAGAGATCTCACCAAGCG	2785
OY	1384	ttaaagcaacttgacctcaaagccatgtatataccaagatatgtgcgtgltttagtgaag	1443

Db	2786	CTCCGATCATGAGTTTTC	CCGATTCATGCTC	TAGACAGAGATGTGAGTTGGTTTATACACCT	2845									
Qy	1444	gaacaacacacaggagcgctt	acacacata	lttagc	atagataagacaaatcaagaa	1503								
Db	2846	AAAAACACCACTGGAGCAT	TTGAC	TACACAGGCTGCCAATGATGCTCTCAACTTTAAAGG	2905									
Qy	1504	gcaacaggttccaggaattg	cgctt	taacaaaat	tgacaactaactatggacttcagtt	1563								
Db	2906	GCTATAGGTTCCAGAGCT	TGGCTGT	ATTATTCACCAATAATATGCAAAATCTTGGGACAGAAATA	2965									
Qy	1564	atcaattcccttataata	tgatgggagat	tgacaattccctgaattc	tgatgtctccagta	1622								
Db	2966	ATTATATCTTCATCAT	TATAGTTGGCACT	TAACTGTACTCTTATGACAAATGTATGCCATC	3022									
Qy	1624	cttgcgcgacaggaat	tgattgaacccg	gaagat	gacatgacatgttccaaagaataag	1683								
Db	3026	ATTGCAATAGCAGAGAT	TTGTTGAAAT	ATAAATGTTGCTGTGGACACACTGAAAGATAG	3083									
Qy	1684	caagaaactaagca	tgctctg	aaagata	tgacaactgaaagcctttggagataaactgata	1743								
Db	3086	AAAGAACTAGAAAGT	GCTGGGAGAGATGCT	ACTAGAAACATATAGAAATCTTCGAAACGTT	3145									
Qy	1744	gtgtcaattacaag	ggaaaaagcc	cttcgagcaaat	gataagaagatgcttcaagctcaa	1803								
Db	3146	GTTTCTTTGACTCTG	AGGACGAGAAATTTGAA	CAATATGATGCTCAGAGTTTGGACGGTACCA	3205									
Qy	1804	caagaaat	lactctgaagaagaacag	atattatggagcgt	gatagcattcgacatgc	1865								
Db	3206	TACAGAAACTCTTTG	AGAAAGCACACAT	TTTTGGAAATTAATTTTCTTCACCAAGGCA	3265									
Qy	1864	ttatataatttgc	ctatgacagcag	ggttcgaatttggagcctat	taattcaagctga	1922								
Db	3266	ATGATGATTTTTTCC	TATGCTGATGTTTCCGGTTTGGAGCCT	ACTTGGTGCCACATATAA	3322									
Qy	1924	cgaatgacc	ccagagagggc	atgattc	atgttaactgtacatgacatgagc	1983								
Db	3326	CTCATGACCTTTG	AGGATGTTCTGTTATGTTTTCACGCTGTCTTCTTGGTCCCATGCGC	3385										
Qy	1984	atcggaaaaaagc	ctcgttttgg	gctccctgaat	cttcaaaagcgaatccggggtctgcgat	2043								
Db	3386	GTGGGGCAATGCT	CACTTCTTATTTGCTCTGACTTATGCCAAACCAAAATATACAGACCCAC	3445										
Qy	2044	ctgttgcctgtt	tggaagaacaa	ccaatatagacagccg	cgatcaagaagaagggaataag	2103								
Db	3446	ATCATCATGATCAT	TTGAAAAA	CCCCCTTGTATTTGACAGCTATACAGCAACGGAAGCCCATAG	3503									
Qy	2104	ccaagacacat	gtlgaaggaa	attagagttc	gagaagctccttctctc	tacatc	atgctgc	2165						
Db	3506	CCGAACACATTTG	AGAAAGAAATGTCACT	TTTGGTGAAGTTGTATTCATCATATCCACCGGA	3565									
Qy	2164	ccaagatgtt	cttccatcc	ccctgcgttga	cttccctcaagat	tgagggaggaagaagcagtagca	2222							
Db	3566	CCGGACATCCCA	AGTCTTCACAGGCACTGAGCTTGAGAGTGAAGAGGCCACGCTGCGCT	3622										
Qy	2224	tttgggggag	caagcggtg	tggaagaagc	actctgtt	ccaacttctgc	agaagc	attat	2283					
Db	3626	CTGTGGGCACT	AGTGTGCTGTGGGAAAGACACAGTGTGTCCAGTCTTGGAAGCGTTTAC	3688										
Qy	2284	gaacccgtg	caagagaca	agtgtgtt	gtatg	atg	atgacaa	gaat	tgaa	atg	2343			
Db	3686	GACCCCTTGGC	AGGAAAGTCTCTCTTGATGACGCAAAAGAAATAAAGCATGAAATGTCAG	3745										
Qy	2344	tgatctcc	gttccccaaat	agcaatc	gtgttcc	ccaagaac	ctgtgtct	cttcaac	gtgcag	gaatt	2403			
Db	3746	TGGCTCCAGACAC	CTTGGGCACTGTGTCCACAGACCCATCTGTTTATGTACGACGACATT	3805										
Qy	2404	gctggaacac	cgagc	ctatgtgt	gacaaagcc	cggtgtgtgc	atla	ataga	lca	aaaga	2463			
Db	3806	GCTAGAACAT	TTGGCTATGTAGAGACAAACAGCGGGTGTGTACACAGGAATAAGATTCCTGAGG	3865										
Qy	2464	gcgcgaat	gtcagcaaat	atc	atc	cttctt	at	tgaa	ag	gtcc	cttga	ata	caaca	2523





Dh 4298 GCTATAGGTTCCAGGCTTGTGTAATTACCCAGATATAGCAAACTTGGACAGAGATA 4357  
Qy 1264 atcattcccttatatatgagtgagatgacatcccgatctcgagatgctccagta 1623  
Dh 4358 ATTATATCTTCATCTATGTTGGCACTACACTGTTACTCTTATGCAATGTCACCCATC 4417  
Qy 1624 ctgcctgagacagatgattgaaacgcgcaatgacatgattgttgcacaagaataga 1683  
Dh 4418 ATTGCATATGACAGAGTTGTTGAAATGAAATGTTGTTGACAGCAAGCACTAAGATAG 4477  
Qy 1684 caagaactaagcagctgagaaagatagcaactgaagccttggagaatatacgtactata 1743  
Dh 4478 AAGAAGTGAAGAGGTGCTGGGAGACATGCTACTGAAGAAATGAAAACCTTCGAACCTT 4537  
Qy 1744 gtgtcattaaacaaggaaagcctcgagcaaatgtatgaaagagatgcttcagaactaa 1803  
Dh 4538 GTTCTTCTTGACTCAGAGCAAGAAAGTTGAACTATGTAATGTAATGTAATGTAATGTA 4597  
Qy 1804 cagagaaatcctgagaaagacagataatgtgaagcgttctgacatgacatgac 1863  
Dh 4598 TACAGAACTTTTGGAGAAAGCAACATCTTGGAAATTAATTAATTAATTAATTAATTA 4657  
Qy 1864 ttatataattgtccatgacagcaggttgcagattgagccttataatcaagctgga 1923  
Dh 4658 ATGATGATTTTCTTCTATGCTGATGTTCCGGTTGGAGCCTACTGTGTGACATATAA 4717  
Qy 1924 cgaatgacccagagagagatgcatagtttctacatgacatgacatgacatgac 1983  
Dh 4718 CTGATGAGCTTTGAGAGATGTTCTGTAATTAATTAATTAATTAATTAATTAATTAATTA 4777  
Qy 1984 atcgaaaaaacgctcgttctgctccctgaatataccaaagcgaacgcgagctcgagat 2043  
Dh 4778 GTGGGGGCAAGCAGTATCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4837  
Qy 2044 ctgttgcctgttggaaagaaacaaatagacagcgcgacgacgacgacgacgacgacg 2103  
Dh 4838 ATCATATGATCATATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4897  
Qy 2104 ccaagacatgltgaaggaattgagattcgagaagctcttcttcttcttcttcttcttct 2163  
Dh 4898 CCGAAGCATATGGAAGAAATGATACATTTGTAAGTTGTAATTAATTAATTAATTAATTA 4957  
Qy 2164 ccaagatgttccatccctcggttccatccctcgatgagcagagagagagagagagagag 2223  
Dh 4958 CCGGACATCCCAAGTCTTCAAGGACTGAGCCTGAGGTGAAGAGGCGCAGACGTGCT 5017  
Qy 2224 ttgttggagagcagcgcgttgggaaagacatctgttcaactcttgcagagacttlat 2283  
Dh 5018 CTGGTGGGACAGTGGCTGTGGGAAGACAGTGTCCAGCTCTGAGCGGTTCTAC 5077  
Qy 2284 gacccctgcaagagagagatgctgttgaatgtgtgagatgcaaaagaattgaatgacag 2343  
Dh 5078 GACCCCTTGGAGGAAAGTCTGCTGATGCAAAAGAAATTAATTAATTAATTAATTAATTA 5137  
Qy 2344 tggctccgttcccaatgacatgcttcccaagagcgttgcgttcaactgacagagact 2403  
Dh 5138 TGGCTCCAGACACACTGCGATCGTGTCCAGAGCCCACTCTTGTGATCGACGACTT 5197  
Qy 2404 gctgagaacatcgctatgtgtgacaacagcgttgtgtcattagatgagatcaaaaga 2463  
Dh 5198 GCTGAGAACTTGGCTTATGAGAGACAGCCGGGTGTGTACAGAGAGAGATCTGTAGG 5257  
Qy 2464 gctgcaaatgcaagaatataatcatcttcttcttgaagcttccctcgagaatatacagca 2523  
Dh 5258 GCAGCAAAAGGAGCCAAATACATCTCTTCAATGAGTCACTGCTTAATTAATTAATTAATTA 5317  
Qy 2524 caagtgtgacgaaagagcagcagcttctgtgcgcgcaagaacaaagcgtcgttca 2583  
Dh 5318 AAGATGAGAGAGAAAGAACTACGCTCTCTGTGGCCAGAAACAGCATGCTCCATACCT 5377  
Qy 2584 agggctcttcccaaaaaccaaatttattgttgaatgagcagcacttcagccctgat 2643  
Dh 5378 CGTGCCCTTGTAGACAGCTCATATTTTGTGTTGATGAAGCAACGCTGAGAT 5437

Qy 2644 aatgacagtgagaagtggttcaagcatgccccttgaataaagcgaagagcgaagacatgc 2703  
Dh 5438 ACAGAAAGTGAAGAGTTGTCCAGAGACCTTGCAAGAGCCAGAAAGCCGACCTGC 5497  
Qy 2704 ctatgtgacatcaagcagctcctgcaatcagaacgcagatlttgaatgtgttcac 2763  
Dh 5498 ATTGTGATTCCTACACCGCTGTCCACATCCAGATGACAGACTTAATGATGTGTTCCAG 5557  
Qy 2764 aatgaaagataaagaaacaaagaaatcaagaagcctcctgagaatcgacatatat 2823  
Dh 5558 AATGCAAGTCAAGAGACATGAGCAGCATCAGCATGCTGTGCGACAGAAAGCATCTAT 5617  
Qy 2824 tttaagttagtgaatgacagatcag 2848  
Dh 5618 TTTTCATATGTCAGTGTCCAGGCTG 5642

RESULT 7  
AR028672 AR028672 9318 bp DNA PAT 29-SEP-1999  
LOCUS  
DEFINITION Sequence 6 from patent US 5858744.  
ACCESSION AR028672  
VERSION AR028672.1 GI:5940645  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9318)  
AUTHORS Baum, C., Stocking-Harbers, C. and Osterlag, W.  
TITLE Retroviral vector hybrids and the use thereof for gene transfer  
JOURNAL Patent: US 5858744-A 6 12-JAN-1999;  
FEATURES  
Source 1..9318  
BASE COUNT 2377 a 2217 c 2364 g 2358 t 2 others  
ORIGIN

Query Match 36.1%; Score 1030.2; DB 6; Length 9318;  
Best Local Similarity 62.2%; Pred. No. 3e-198;  
Matches 1695; Conservative 0; Mismatches 980; Indels 42; Gaps 3;

Qy 166 attatgtaagaacccaagatatagatacttcttccacagcgtgatataaactaatcc 225  
Dh 2877 ATATTTGATTAATTAAGCCAGATTTGACAGCTATTGGAAGAGTGGCACAAACCATATAT 2936  
Qy 226 atagaaggaactgtgaaatttaaaatgttcttcaatatacacaagacacatc 285  
Dh 2937 ATTAAGGGAATTTTGGAATTCAGAAATGTTCACTTCACTTACCTTACCTTACCTTACCTT 2996  
Qy 286 aagattctgaagagtcgaatcagaatcagaatgaagtcgagagagacagtcgctgtgcgt 345  
Dh 2997 AAGATCTTGAAGGCGCTCAACCTGAAAGTGCAGAGTGGGACAGACGCTGCTGTTTGA 3056  
Qy 346 ctcaatgacagtgagaaagagatagcgttgcagcttgcagagagttatagatccgat 405  
Dh 3057 AACAGTGTCTGTGGGAAGGCAACAGTCCAGCTGATGACAGGCTCTATGACCCACA 3116  
Qy 406 gatgcttcatatgttgaatgagaatgacatcaagcgtttaaagtgctgcgcatatcga 465  
Dh 3117 GAGGGGATGTCAGTGTGATGAGACAGATATTAGACCATTAATGTAAGTTTCTTACGG 3176  
Qy 466 gaccatattgagatggttagtcaagaagcgttctgttcgaggaaccaatcagtaaat 525  
Dh 3177 GAATTCATTTGTTGTGTGATGAGCAACCTGTATTGTTCCACACATATACCTAATAAC 3236  
Qy 526 atcaagatgacagagatgagatgacatgagagagatgagagagagagagagagagag 585  
Dh 3237 ATTGCTATGGCCGTGAAGAAATGTCACATGATGATGATGATGATGATGATGATGATGATG 3296  
Qy 586 aatgctatgatttctatcatgagatgttcttaataatlaacatctgtgaggggaaaaa 645

Dh 3297 AATGCTATGACTTTATCATGAACTGCTCATTAATTTGACACCCCTGGTTGGAGAGA 3356  
Oy 646 ggaagcgaatgagtgagggcgaacacagagatgcgaattgctgcttgcttaagtcga 705  
Dh 3357 GGGGCCAGATTGAGTGGTGGGACAAAGAGATGCCCATTTGACAGTGGCCCTGGTTCCG 3416  
Oy 706 aaccccaagatcttgattatagatgagctacgctgcgcccgttgatcagaagaagtcga 765  
Dh 3417 AACCCCAAGATCCCTGCTGAGTGAAGGACGTCACCCCTTGGACACAGAAAGGAAGCA 3476  
Oy 766 gctgttaagctgcagtcggaagaagcgagcaaaagtcgagctaacacgctggtgagcaac 825  
Dh 3477 GTGGGTACGTTGAGCTCTGGATTAAGGCCAGAAAGGTGGAGACACCATTTGATAGCTCAT 3536  
Oy 826 cgaattctactactcgaagtgcagatttgatgtgacccctaaaggaaggaaatgcgagc 885  
Dh 3537 CGTTGCTACAGTTGCTGAATAGCTGACGTATCCCTGGTTTGCATGATGGAGCATTTGTG 3596  
Oy 886 gagaagaagacacatgctgaacataaggcaaacgaggtctatatattcaacttgatg 945  
Dh 3597 GAGAAAGGAATCATGATGAACCTCATGAAGAGAAAGGCATTTACTTCAAACTTGTCACA 3656  
Oy 946 tcaagagatatlaaaaaagctgagacagatgagtcgaatgacatatctca----- 997  
Dh 3657 ATGCAAGACGACGAGAAATGAAGTTGAATTAAGAAATGACGCTGATGATCCAAAAGTGAA 3716  
Oy 998 -----ctgaagaagaagccaactcaactcctctgcagctcgtgaagagcaca--- 1045  
Dh 3717 ATTATGCTCTTGAATGTCTTCAAAATGATTCAGATCCAGTCTAATTAAGAAAAAGATCA 3776  
Oy 1046 -----agtcagactcattgacaagctgaggaatcccaaccaactc 1086  
Dh 3777 ACTGATGAGAGTGTCCGCTGATCAAGACCCAGACAGAAAGCTAGTACCAAGAGAGCT 3836  
Oy 1087 aaagaagataagcttctctgaaagctcctctatlaaaaaattlaaagttlaaaaaagcagaa 1146  
Dh 3837 CTGATGAATATATACCTCCAGTTCTTTGGAGGATTAAGAGCTTAATTAATTAATTAATTA 3896  
Oy 1147 tggcctcttgctgcttgaggaacatgagctcgtctctcaatggaacatgctcaccagta 1206  
Dh 3897 TGGCTTATTTTGTGTTGGTGTGCTATTTTGTGCCATTAATTAATTAATTAATTAATTAAT 3956  
Oy 1207 tttccatcattcttgcaaaaaataacacatglttggaaataagataaaac-----aca 1263  
Dh 3957 TTTTCATAATTAATTTTCAAAATTAAGGGTTTTACAAATTAATTAATTAATTAATTAAT 4016  
Oy 1264 ttaagaagctgcaagaatattatccatgattctgcattggtggttaattgctt 1323  
Dh 4017 AAACGACAGAAATGAACTTGTCTTCACTATTGTTTCTAGCCCTTGGAAATTAATTTCTTT 4076  
Oy 1324 gtcagttattcatcagcagagatttlaacgcagagcagaggaagaattlaacgatagaga 1383  
Dh 4077 ATTACATTTTTCCTTCAGAGGTTTCACATTTGGCAAGCTGAGAGATCTCCACCAAGGG 4136  
Oy 1384 ttaagaacatgtgccttcaaaagcattgtatatacaagattatgcctggttgatgaaag 1443  
Dh 4137 CTCGATACATGATGTTTCCGATTCATGCTCAGACAGATGATGATGATGATGATGATGAT 4196  
Oy 1444 gaaacacgcagagagccttgacaacaatalatagccatagatatagcaacaatcaagga 1503  
Dh 4197 AAAAGACACACTGAGAGCTTACTACACAGGCTCGCCATTAATGCTGCTCAAGTTAAAGG 4256  
Oy 1504 gcaacagttcagagatggggtcttaacacaaaatgacaatacaatcaatgagccttcagt 1563  
Dh 4257 GCTATAGTTTCAGGCTTGCCTGTAATTAACCCAGAAATTAAGCAAAATCTTGGACAGGAATA 4316  
Oy 1564 atcattctctatataatgagtgagagatgacatccgattctgagatgagctccagta 1623  
Dh 4317 ATTATATCTTCATCTATGTTGGCACTAACAAGCTGATCTTATAGCAATTTATGCCCATC 4376  
Oy 1624 ctgctgcgtgagagaaatgaaacgcgcaatgactgagatttgccaacaagaatag 1683  
Dh 4377 ATTGCATATGACGAGTGTGAAATGAATAATGTTGTCTGACAAAGCACTGAAAGATTAAG 4436

Oy 1684 caagaactlaagcatgcttgaaagatagcaactgaagccttggagatatagctactata 1743  
Dh 4437 AAAGAACTAAGAGTGTGGGAAAGATCGTACTGAAGCAATTAAGAAACTCCAAACCGTT 4496  
Oy 1744 gttctatlaacaagggaaaaagccttgacgaatgatatgaagatgtgctgactacaa 1803  
Dh 4497 GTTCTTTGACTCAGGAGCAAGATTGAACATATGATGATCTCAGATTTGCAAGTACCA 4556  
Oy 1804 cacaagaataccctcgaagaagaagcagattatggaacggttatagcatcaagcattgccc 1863  
Dh 4557 TACGAAACCTTTGAGGAAGACACACATCTTTGGAATTAATTTCTTCCACCAAGCA 4616  
Oy 1864 ttatatattctgctatgcaagcaggttccgatttgagcttgaacatttaactaacgtgga 1923  
Dh 4617 ATGATGATATTTTCCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4676  
Oy 1924 cgaatgccccagagggcagtgatgatttcttactcgaattgacatgagatgagc 1983  
Dh 4677 CTCTAGAGCTTTGAGGATGTTCTGTTAGTATTTTCACTGTTTGTGTTGTTGTTGTTGTT 4736  
Oy 1984 atcgaaaaaagcctgcttctgctccgaatatccaaagccaatcgggggtcgcgcat 2043  
Dh 4737 GTGGGGCAAGTCAGTTGATTTGCTCTGACTATGCCAAAGCCAAATATCAGACGCCAC 4796  
Oy 2044 ctglttgcttgttgaaagaacccaatatagacagccgagltcaagaaggaaag 2103  
Dh 4797 ATCATCATGATCATTTGAAAAAACCCTTTGATTAACAGTACAGACGAGAAAGGCTTAATG 4856  
Oy 2104 ccagaacatgtaaggaattatagatltcgaagaatccttctctctctctctctctctcgc 2163  
Dh 4857 CCGAACAATTTGAGGAAGAAATGATCATTTGGTGAATTTGATTTGATTTGATTTGATTTG 4916  
Oy 2164 ccagatgtttcatcctccgctgcttccctcgaagtatgagcaggaagaacagatga 2223  
Dh 4917 CCGACATCCAGTGTCTTCAAGGACTGAGCTGAGAGTGAAGAGGCCACAGCCCTGGCT 4976  
Oy 2224 ttgtgggagcagcagcgtgltggaaaaagcactcgttcaactctgcagagacttat 2283  
Dh 4977 CTGCTGGGACAGCAGTGGCTGTGGGAAGACACAGTGTGCTCCAGCTCTCGAGCGGTTTAC 5036  
Oy 2284 gaccccgtaaggaagaagctgttgatggtgtgtagatgtagaagaagaatgtagatga 2343  
Dh 5037 GACCCCTTGGCAGGAAAGTGTCTGTGATGAGCAAAAGATTAAGCACTGAATGTTACG 5096  
Oy 2344 tgcctcgctcccaatagacatcgtctcccaagaagcgtgcttcccaactgagcatt 2403  
Dh 5097 TGGCTCCAGACACACCTGGGATGCTGTGCCAGAGGCCCATCTGTTGATCGCAGCAT 5156  
Oy 2404 gctgagaacalcgctatgtagcaacagccgltgtgtgcattagatagatcaagaag 2463  
Dh 5157 GCTGAGAACTTGGCTTATGAGACACAGCCGGTGTGTACAGGAAGAGATGTTGAGG 5216  
Oy 2464 gccgcaatgagcaaatatcattcttattgaagctccctcgtggaataaacaaga 2523  
Dh 5217 GCACCAAGGAGGCAACATATGCTCTTCAATGATGATGATGATGATGATGATGATGATGAT 5276  
Oy 2524 caagtggactgaagagcagccttctggtggcgaagaacaaagactgattgca 2583  
Dh 5277 AAAGTAGAGACAAAGACACTCAGCTCTGTGTGGCCAGAAACACGATTTGCCATGACT 5336  
Oy 2584 agggctctctccaaaacccaataattlatgttgatgagcgaactcagccctcag 2643  
Dh 5337 CGTGCCCTTGTAGACAGCCCATATTTTGTGTTGGATGAAGCCAGCTCAGCTCGAT 5396  
Oy 2644 aatgacagtgagaaggtgtgtcagcagtccttgatgaagccaggaagcaggaagacatgc 2703  
Dh 5397 ACAGAAAGTGAAGGTTGTGTCCAAAGAGCCCTGGACAAAGAGAGAGGCCGACCTGC 5456  
Oy 2704 ctgattgtcactcaagcgtctcgcgaattcagaacgcagattgtagatgtgttcgac 2763  
Dh 5457 ATTGATTTGCTCACCCGCTTCCACCATCCAGATATGCTACAGCTTAATATAGTGTGTTACG 5516

OY	2764	aatggaagataaaaggacaagaactcatcaagcctcccttgaagaatcgagaacatatat	2823
Dd	5517	AATGGCAGAGTCATAACGACATCGGCACGATCGACACTGTGGCACAGAAAAGCATCTAT	5576
OY	2824	ttaagttagtcgaatgcacagtcag	2848
Dd	5577	TTTTCAATTGCTCAGTGCTCCAGGCTG	5601
RESULT	8		
LOCUS	AR051647	4264 bp	DNA
DEFINITION	Sequence 1 from patent US 5830697.	PAT	29-SEP-1999
ACCESSION	AR051647		
VERSION	AR051647.1	GI:5975011	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified. 1 (bases 1 to 4264) Sikic,B.I. and Chen,G. P-lycoprotein mutant resistant to cyclosporin modulation Patent: US 5830697-A 1 03-NOV-1998; Location/Qualifiers source 1..4264 /organism="unknown"		
BASE COUNT	1263 a	807 c	1041 g 1153 t
ORIGIN			
Query Match	36.0%;	Score 1028.6;	DB 6; Length 4264;
Best Local Similarity	62.2%;	Pred. No. 6.4e-198;	
Matches 1694;	Conservative	0; Mismatches 989;	Indels 42; Gaps 3;
OY	166	attatgtataaagaaccggatagataaactttcacagctgatataacctgaatcc	225
Dd	1240	ATAATTGATTAATAAAGCCAAGATTTGACAGCTATTGGAAGATGGGCCACAACCAAGATAAT	1299
OY	226	atagaaggaaactgttgaaactttaaaaatgtcttccttcaattaccatacaagacatctatc	285
Dd	1300	ATTAAAGGGAATTTTGGAAATTCGAATATGTTCACTTCACTTACCATTCTCGAAAAAGAGTT	1359
OY	286	aagattctyaaagtcctgaatctcagaattaaagtccttgagaagacaagctgccttgcgt	345
Dd	1360	AAGATCTTGAAGAAGCGCCTGAACCTGAAGGTGCAGAGTGGGACAGACGTCGTCCTGGTTGGA	1419
OY	346	cctaattgcgaagtggaagaagatgcgttagtcagcgtcttcgcagaagttatatgcctgat	405
Dd	1420	AACAGTGGTGTGGGAAGGACCAACAGTCCAGCTGATGACAAAGGCTCTATTATACCCACA	1479
OY	406	gattgcttatcatgctgtgatgatagaatgacacacagagctttaatgtygcgcatatcga	465
Dd	1480	GAGGGGATGTCAGTGTTGATGATGACAGAGATATTAGACACCATTAATGTAAGTTTCTACGG	1539
OY	466	gaccalattggagtggttagtcaagaagcctgtlittgltcggagccaacatcagtaacaat	525
Dd	1540	GAAATCATTTGCTGTGTGATGATGACGAACCTGTATTGTTTCCACACACGATACCTGAANAAC	1599
OY	526	atcaaagttagcgcgagatgatgttgtactgaatgaagaagatvgbaaagcgcaaggagaaga	585
Dd	1600	ATTTCCTAATGGCCGGGAAATATTCACCATGATGATGAGATTGAAGAAAGCTTCAAGGAAGCC	1659
OY	586	aatgctaatgattctatcacctggagttcccttaataattatatacttggtagggaaaaa	645
Dd	1660	AATGCTTAAGATTTTATCTATCGAAACTGCCTCTAATAATTTTGACACCCCTGGTTTGGAGAGAA	1719
OY	646	ggagctcaaatgagtggagggcagaaaacagagatcgcaatgtcgtcgttacttactga	705
Dd	1720	GGGGCCCACTTAGTGTGGTGGCGACAGAACGAGAGATGCCCATTTGACAGTCCCTGGTTGCC	1779
OY	706	aaccccagaatctcyatttttagatgaagctacgctctgccctgagattcaagaagaagta	765
Dd	1780	AACCCCAAGATCTCTCTGTGTGATGATGGCCAGTCAAGCTTGTGACACAGAAAAGCAATCAT	1839

QY	766	gctgttcaagctgcacgtcgagaaagcgagcaaaagtcgcgaactcaacatgltgtgtacac	825
Db	1840	GTGGTTCAGGGCGCTCGTGATTAAGGCCAGAAAAAGGCGGAGCACCACTGTGTATACCTCAT	1899
QY	826	cgacttctactatctgaaatgcagaatttgatgttgaccctaaagatgaaatgcgtggc	885
Db	1900	CGTTTGTCACAGTTTCCTAATGCTGACGTCATCGCTGGTTTGATGATGAGACTCATTTGTG	1959
QY	886	gagaagaagacacatgctgcacaaatgycaaaacgagctcatatcatcacttgatg	945
Db	1960	GAGAAAGCAATCATGTATGTAACCTCATGAAGAAAGGCAATTTACTCTCAAATCTGTACA	2019
QY	946	tcacaggaatataaaaaagctgatagcagatgtagtcaatgacatattctca-----	997
Db	2020	ATGCGACAGCAGGAAATGAAAGTTGAATAAGAAAGACGCTGATCAATCCAAAAGTGA	2079
QY	998	-----ctgaagaagaacacacactcactctctcctgcactctgtaagacalca---	1045
Db	2080	ATTGATGCGCTTGGAAATGTCTCTCAATGATTCAGATCCAGTCCTAATGAAGAAAAGATCA	2139
QY	1046	-----agtcagactcatitgacaagagctgaggaatccaccatct	1086
Db	2140	ACTCGTAGAGTGTCCGTGATGATCAAGGCCCAAGACAGAAAGCCTTAGTCCMAAGAGCT	2199
QY	1087	aaagagataaagctctctgtaagctctctcatltaaaaatttaagttaaacagaactgaa	1146
Db	2200	CTGGATGAAATATTAACCTCCAGTGTCTCTTTGGAGCATTTATGAAGCTAATTAATTA	2259
QY	1147	tggccttttggtgtctcgggagacatggtctctgtcttaaatggaactgtcatcagta	1206
Db	2260	TGGCTTATTTGTGTGTGGTGATATTGTGGCATTTATAAATGAGAGCCTGCACACACCA	2319
QY	1207	tttccatcalcctltgcaaaaatataacaaatglttggaataatgataaaac---aca	1263
Db	2320	TTTGCATAATTAATTTTCAAAGATATATAGGGGTTTTTACAAAGATTCATGATCCTGMAACA	2379
QY	1264	ttaagcagltgcaagaatttatccatgabatltgcataatttggtgtgtatgtgctt	1323
Db	2380	AAACGACGAATTAACCTGTTGTTTACATATTTGTTCTGACCCCTTGGAATTAATTTCTTTT	2439
QY	1324	gtcagttatcttcacgacaggaattatlttaacgacagacagggaaatttlaacgaatga	1383
Db	2440	ATTACATTTTTCCTTCAAGGGTTTCACTTTGGCAAAAGCTGGAGAGATCCTCACCAAGCG	2499
QY	1384	ttaagacacttggccttccaaagccaaatgatatccagagatatgtgctgtgttgataaag	1443
Db	2500	CTCCGATTCATAGGTTTTCCGATCCATCCTCCAGACAGAGATGTAGTGGTTTATACACCT	2559
QY	1444	gaaacagacagagagctgttgcacaaatatatgacatatagataagcaaaatcaaaag	1503
Db	2560	AAAAACACCACTGGACGATTTGACTACCAAGGCTGCCAATGATGCTGCTCAAGTTAAAGG	2619
QY	1504	gcaacaggttccagagattggcgtcttacaacaaatgcaactaacaatggaacttcagtc	1563
Db	2620	GCTATAGGTTCCAGGCTTGCTGTATTAATACCCAAATATATGCAAAATCTTGGACAGATTA	2679
QY	1564	atcatcttccattatatatgatatgtagatgagatgacatltccgatltccgatattgtccagta	1623
Db	2680	ATTATATCTTCATCATATGTTGGCACTTAACACTGTCTCTTAGCAATTTTATCCCATTC	2739
QY	1624	cttgcgcgtaagaaatgatatgtaaacgcgcaatgacatgtgatttccacaagaataag	1683
Db	2740	ATTGCAATAGCAGAGGTTGTTGAATATAAAATGTGTGCTGACAGCACTGAAATAATAG	2799
QY	1684	caaggaacttaagcaltgctgtaaaatagcaactgaaagcttltggagataatacgtactata	1743
Db	2800	AAAGACTAGTAAGAGTGCTGGGAAGATCGTACTGAGACAAATAGAAAACTTCGAGACGGT	2859
QY	1744	gtgtcatataacaagggaaaaagccttcogagcaaatgatatgaaagagatgtcttaagctcaa	1803
Db	2860	GTTTCTTTTGACTCGAGGAGCAAGAAATTTGAACAATATGTATGCTTCAGAGTTTGCAGGTACCA	2919



OY	886	gaaagaagacacacitctgaacctaagcgaaaagagcgtatataatcactgtgag	945
Db	1360	GAGAAAGGAATCATGATGATGACATGAAAGGAAAGGAGATTACTTCAACTGTGCACA	2019
OY	946	tcaacagatataaaagcttgaatgacagatgaggtcactgacatgacatctca-----	997
Db	2020	ATGCAAGACACAGAGAAATGAAGTTGAAATTAGAAATTCGACGTGATGATCCAAAAGTGA	2079
OY	998	-----ctgaagaagaacccaactcactcctctgcactcttgaagagcatca--	1045
Db	2080	ATTGATGCTTGGAAATGCTTTCAAATGATTCAGATCCAGCTATTAAGAAAAAGATCA	2139
OY	1046	-----agtaagcttcattgnaacagagcctgaggaatccaccaatct	1086
Db	2140	ACTCTGATGAGATGCTCCGTGATCATCAAGCCCAAGACGAAAGCTTAGTACCAAGAGCT	2199
OY	1087	aagaagaatagctctcctgaagctctcattcaaaatttaagttaaacaagcgtga	1146
Db	2200	CTGATTAAGTAAGTATACCTCCAGTTCCTTTTGGAGAGATTATACCTAAATTTAACTGA	2259
OY	1147	tggccttttggcttctgagggacatgctctgtcttaaatgaaactgttcatcagta	1206
Db	2260	TGGCTTATTTTGTGTGGTATTTTGTGCATTAATTAAGAGGCTGCACCAACACA	2319
OY	1207	ttttccatcctcttgcaaaaataataccagtgttggaaataatgataaaac---aca	1263
Db	2320	TTTGCAATTAATATTTTCAAGATTAATAGGGGTTTATACAAAGAAATGATGATCCGAAACA	2379
OY	1264	ttaagaatgatgcagaaaatttatccagatattgcattcttgggtgattatgcttt	1323
Db	2380	AAACGACGAATAGTAACTTGTTTACTATATTGTTCTAGCCCTTGGAATTAATTTCTTTT	2439
OY	1324	gtcagtatattctgcaaggaattattctacggcagagcaggggaaattttaacgtaga	1383
Db	2440	ATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAGCTGGAGAGATCCTCACCAAGCG	2499
OY	1384	ttaagaacttggccttcaaaagccatgtatatacaggaatatgctgtgttgaagaag	1443
Db	2500	CTCCGATACATGATGTTTCCGATCCATCAAGCTCAGACAGAGATGATGATGTTGATATACCT	2559
OY	1444	gaaacacgcacagagagcttgaacaacatattagccatagatatagcaacaatcaaga	1503
Db	2560	AAAAACACCACTGAGCATTTGACTACCAAGCTCCCAATGATGCTGCTCAAGTTAAAGG	2619
OY	1504	gcaacaggttccaggaatttggcgtcttaacaacaatgacaactcaaatggaacttcagt	1563
Db	2620	GCTATAGGTTCCACAGGCTTGCTGTATTTACCCAGAAATATGCAAACTTTGGACAGAGATA	2679
OY	1564	atcattccttataatatgatatggatgagatgacatccctgattctgagtattgtccagta	1623
Db	2680	ATTATATCTTCACTATAGTTGTCGAACATACACGTTTCTCTTGACAAATGTATACCAATC	2739
OY	1624	cttgcggtgcagaagaatgattgaaccgcgacaatgacatgagatttggcaacaagaatag	1683
Db	2740	ATTTCGAATACAGGAGTTGTTGAAATGAAATGTTGTCTGCAAGCAACCTGAAAGATTAAG	2799
OY	1684	caagaacttaagatcgtctgnaaagatagcaactgaaacttggagaatataagtcata	1743
Db	2800	AAAGAACTAGAAAGTGCTGGGAAGATTCGCTACTCGAAGCAATAGAAACTTCGAAACGTT	2859
OY	1744	gtgtcattaaacaagggaaaagccttcgcagcaaatgatalgaagagatgctcagaatcaa	1803
Db	2860	GTTTCTTTGACTCAGAGCGAGAGTTTGAACATATGTATGCTCAAGATTTGCAGGTGACCA	2919
OY	1804	cacggaataaccgcgaagaagaacagatattatggaagctgtatgaacttaagccatgcc	1866
Db	2920	TACAGAAATCTTTGAGAAAGACACATCTTTGGAAATTACATTTTCTTCACCCAGCA	2979
OY	1864	tttatataatttgcctatcagcagaggttctgatttggagcctatataatcaagctga	1923
Db	2980	ATGATGATATTTTTCATATGCTGATGTTTCCGTTTGGAGCCTACTGTGTCACATATAA	3039
OY	1924	cgaatgaccccgagggcagtgatcatagtttctaactgcaattgacatabgagctatggcc	1983

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
E02326	Multi drug resistance relating gene derived from human normal cells	E02326	4378 bp	RNA	PAT 29-SEP-1997
Db	3040	CTCATGACCTTTGAGGATGTTCTGTTAGTATTTTCAAGCTGTTGCTCTTGGTCCATGGCC	3099		
Qy	1984	atcggtaaaacgcctcgttttttgctctctgaataatccaaagccaaatcggggctgcgcat	2043		
Db	3100	GTGGGGCAAGTCAGTTCACTTTTGCCTCTGACTATTTGGCAAAAGCCAAATATACAGACGCCAC	3159		
Qy	2044	ctgttttccttgcttgtaaaagaacccaatatagacacgcgcagtcacgaagaagggaaaag	2103		
Db	3160	ATCATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTACAGCATGGAAGGCTTAATG	3219		
Qy	2104	ccaagcacatgtagaaggaaatttagaatttcgaagaagctcttcttcacatcgaatgcgc	2163		
Db	3220	CCGACACATTTGGAAAGGAAATGTCACTTTTGGTGAAGTTGTAFTTCAACTATCCACCCCA	3279		
Qy	2164	ccaagatgttttcatctccctgcgttgcttaaccctcagtatltgaacggaagaaacagtaaga	2223		
Db	3280	CCGGACATCCCAAGTCTCTTACGGAGACTGAGSCTGGAGGTGAANAAGGCCACAGCGTGCT	3339		
Qy	2224	tttgttggaagcagcgcgctgttggaagaacacactctcgttcaactctgcagaagcttat	2283		
Db	3340	CTGGTGGGCGAGCAGTGGCTGTGGGAAGAGCACAGTGTGTCACACTCTCTGGAGCGATTCTAC	3399		
Qy	2284	gaccccgtagcgaagcgaagtgtctgttttagatgtgtgtgaagcgaagaagattgaattacag	2343		
Db	3400	GACCCCTTTGGCGAGGAAAGTGTCTGTTGATGGCAAAAGAAATTAAGCGATGTAATGTTCAG	3459		
Qy	2344	tggtctccgtttcccaatatagcaatcgttctcccaagaagcctgtgctcttcaactgcagcat	2403		
Db	3460	TGGCTCCAGCAGCACACTTGCGGCTGCTGTGTCGCCAGAGGCCATCTGTTTGACTCGACGACT	3519		
Qy	2404	gctgaagaatcgcgctatgtgtgatacaacagcgcgtgtgtgtgacatagatgatatcaagaag	2463		
Db	3520	GCTGTGAACATTTGCTTATGAGAACAAACAGCCGGGTGGTGTCAAGGAGAGATTGTGAGG	3579		
Qy	2464	gccgcaaatgtcagcaaatatcatctctttatgtgaagtgctccctcgagaatatcaaacaca	2523		
Db	3580	GCAGCAAAAGGAGCCCAACATACATGCTCTTCATGAGTCACTGCTTAATTAATATAGCACT	3639		
Qy	2524	caagtttgacgtgaagaagcagacgtttctgcgcgcgaagaacaagaagactgaactatga	2583		
Db	3640	AAAGTAGGAGAGAAAGAAAGAACTCAGCTCTCTGTGGGCCAANAACAGCTTGCCATAGCT	3699		
Qy	2584	agggtctctcttccaaaaacccaataattatgtttgagatgagccacttcagccctcgat	2643		
Db	3700	CGTGCCCTGTTGATGACAGCCTCATATTGCTTTGGATGAAAGCCACGTCACACTCTGAGAT	3759		
Qy	2644	aatgcacgttgagaagatgtgtttcagatgcgccttgataaagccaagacggaagagacatgc	2703		
Db	3760	ACAGAAAGTGAAGAAAGTTGTCCAAAGAGCCCTGGACAAAGCCACAGAGAAAGCCGCACTTC	3819		
Qy	2704	ctagtggtcaactcacaagcgtctctctgcaattctgaagaacgacgaattgataatgttctgcac	2763		
Db	3820	ATTGTGATTTGCTCACCGCTCTGCCACATCCGAATGCGAATTAATAGTGGTGTTCAG	3879		
Qy	2764	aatggaagaatgaagaagaacaagaaactcacaagaagctctcttgagaatcagaacatatat	2823		
Db	3880	AATGCACAGGTCAAGAGAGATGACACGCATTCAGACGCTGCTGGCACAGAAAGCAATCTAT	3939		
Qy	2824	tttaagttagtgaaatgacatgcacgtcag	2848		
Db	3940	TTTTCAATGTCAGTGTCCAGGCTG	3964		



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 4378)  
 Ueda, K. and Komano, T.  
 HUMAN NORMAL CELL-DERIVED MDR RELATED GENE  
 Patient: JP 1990100680-A 1 12-APR-1990.  
 JOURNAL  
 SUNTORY LTD  
 OS Homo sapiens  
 PN JP 1990100680-A/1  
 PD 12-APR-1990  
 PF 05-OCT-1988 JP 1988251475  
 PI UEDA KAZUMITSU, KOMANO TORU  
 PC C12N15/12,C12N1/21,C12O1/68;  
 CC strandedness: Single;  
 CC topology: Linear;  
 CC \*source: tissue: Adrenal gland;  
 CC \*source: clone-SAM132;  
 FH Key Location/Qualifiers  
 FT 3'UTR Location/Qualifiers  
 FT 1..4378  
 FT /organism="Homo sapiens"  
 FT /db\_xref="taxon:9606"  
 BASE COUNT 1315 a 818 c 1057 g 1188 t  
 ORIGIN  
 Query Match 36.0%; Score 1028.6; DB 6; Length 4378;  
 Best Local Similarity 62.2%; Pred. No. 6.4e-198;  
 Matches 1694; Conservative 0; Mismatches 989; Indels 42; Gaps 3;  
 QY 166 attatigataagaaccaglatagataactttccacagctgataaaactgaatcc 225  
 DB 1239 ATATATGATATATAGCCAGATATGACAGTATTCGAAGAGTGGCCACCAAGATAT 1298  
 QY 226 atgaagaagactgtgaattaaatgttcttcaatcatcatcaagaccatc 285  
 DB 1299 ATTAAGGGAATTTGGATTCAGAAATGTTCACTTCAGTTACCATCTCGAAAGAAATT 1358  
 QY 286 aagatctgaagaagctgaatctcaagaatcaagtcgagagacagctgcgctgtcg 345  
 DB 1359 AAGTCTTGAAGGCTCGAACCTGAAGGTGACAGTGGGACAGCGTGGCCCTGGTTGGA 1418  
 QY 346 ctcaatgacagtggaagagtagcagtagtcaagctctcgcagaggtlatalatgcgat 405  
 DB 1419 AACGCTGCTGTGGAGAGACACAGTCCAGCTGATGACAGCTATGACCCCA 1478  
 QY 406 gatgacttcatcagtgatgagatgaatgaatcagagcttaaatgtgcgcatc 465  
 DB 1479 GAGGGAGTGTGCTGATGAGACAGGATATTAGACCAATTAAGGTTCTACGG 1538  
 QY 466 gaccatctgagagtgatgacgaagagctgttcttcgagaccacatcaagaa 525  
 DB 1539 GAATTCATGCTGTGGTGAAGACCTGTATTGTTTCCACCATGATAGCTGA 1598  
 QY 526 atcaagatagcagagatgatgactgatatgaagagatggaagcagcagaagaa 585  
 DB 1599 ATTGGCTATGGCCGTGAATAATGTCACCATGATGATGGAAGAGCTGCAAGAGAACCC 1658  
 QY 586 aatgcgatgatttcatcagaggttccctaaatcaatcaatcagtgatgagagaa 645  
 DB 1659 AATGCTATGACTTATCATGAATACTGCTCATTAATTTGACACCCCTGGTGGAGAGA 1718  
 QY 646 ggaagctcaaatgagtgagagcagaagagagatgcgaatgtcgtgcttgatcga 705  
 DB 1719 GGGGCCAGTTGAGTGGGAGAGAGAGATGCCATTGACGTCCTGGTCTGT 1778

QY 706 aaccocaagattcattttagatgaggtacgtctgccttgattcagaagaagcaag 765  
 DB 1779 AACCCCAAGATCTCTCTGATGAGGCGACGTCACCTTGACACACAAAGCAGCA 1838  
 QY 766 gctgttcaagctgcaacttggaagagcgaagcgaagctgactacaatcgttgacacac 825  
 DB 1839 GTGCTTCAGGTGCTCTGATTAAGGCGCAAAAAGTCGACCACTGTGATAGCTCAT 1898  
 QY 826 cgacttctactacttcgaagtgcagattgattgtgacccctaaagatgagatgctgcg 885  
 DB 1899 CGTTGCTTCAAGTGTGATGCTGACGTCATCGCTGTTTGATGATGATGATGATG 1958  
 QY 886 gagaagaagcagactgctgaactaagcgaagcgaagcgaagcgaagcgaagcgaag 945  
 DB 1959 GAGAAAGGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2018  
 QY 946 tcaagagatataaagaagctgacagatgagtgagtgatgacatctca----- 997  
 DB 2019 ATCCAGACAGCAGGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2078  
 QY 998 -----ctgaagaagaagcgaactcaactcctcctgacactcgtgagagacata--- 1045  
 DB 2079 ATTGATGCCCTTGGAATGCTCTCAATGATTCAGATCCAGTCAATTAAGAAAGATCA 2138  
 QY 1046 -----agtcagacttcattgacaagcgtgaggaatccaccacatct 1086  
 DB 2139 ACTGTAGAGTCCGCTGATCAGACCCCAAGACAGAAAGCTTGTACCAAGAGCT 2198  
 QY 1087 aagaagataagctctccctgagctcctcctataaatttaagtaagtaagcgtgaa 1146  
 DB 2199 CTGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2258  
 QY 1147 tggccttctgtgtctgtgagacatgagctcctcctcctaaatgagacgtgtcctcagta 1206  
 DB 2259 TGGCTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2318  
 QY 1207 ttccatcactcttgcaaaattataacatgtttggaatgataaagc---aca 1263  
 DB 2319 TTGCAATTAATTTTCAAGATTTATGAGGTTTTCACAAATGATGATGATGATGATGATGAT 2378  
 QY 1264 ttaagcagatgcaagaatttatacagatcagatcagatcagatcagatcagatcagat 1323  
 DB 2379 AAGCGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2438  
 QY 1324 gtaagttatcagcagagatatttcaagcagcagcagcagcagcagcagcagcagcag 1383  
 DB 2439 ATTACATTTTCTTCCATGAGGTTTCCATTTGCGCAAACTGAGAGATCCACCAAGCGG 2498  
 QY 1384 ttaagcagcttgctcctcaagcagcagcagcagcagcagcagcagcagcagcagcag 1443  
 DB 2499 CTCGATACATGTTTCCGATCCATGCTCAGACAGATGATGATGATGATGATGATGATGAT 2558  
 QY 1444 gaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1503  
 DB 2559 AAAAACCACACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2618  
 QY 1504 gcaacagcttccagagatgagcgttcaacaacaaatgcaacaaatgcaacaaatgcaacaa 1563  
 DB 2619 GCTATAGTTCCAGGCTGCTGATTAATACCCAGAAATTAAGCAAAATCTTGGACAGGATA 2678  
 QY 1564 atcattccttataatagatgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1623  
 DB 2679 ATTATATCCTTCACTATGTTGTCATCACTAATCACTGTTACTCTTAAGCAATTTGTACCATC 2738  
 QY 1624 ctgacgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1683  
 DB 2739 ATTGCATAGCAGAGATGTTGATGAATGAATGTTGCTGCAACAAGCAGCAAGAGATGAG 2798  
 QY 1684 caagacttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1743  
 DB 2799 AAGAACTAGAAAGTTTGGAGAGATCGTACTGAAGCAATAGAAACTTCCGAACGCTT 2858

```

OY 1744 gtcattacaagaagaaagcccttcgacgaatgataagagatgcttcagactca 1803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2859 GTTCTTTGACCCAGAGAGAGAGATTGACATATGATGCTCAGACAGTTGACAGGTACCA 2918
OY 1804 cagagaatcctcgaagaagaacacagatltatgaagctgtlatgcatcagcatgccc 1863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2919 TACAGAAACTCTTTGAGAAAGCACACATCTTTGGAATATACATTTTCCCTTCCACGAGCA 2978
OY 1864 ttataatatttgctatgacagcaggttctgatttggagcctatttaactgaactgga 1923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2979 ATGAGTATTTTCTTATGCTGAGATGTTCCGGTTGGGCGCTACTGTGCGCATATAA 3038
OY 1924 cgaatgacccacagagcagctgttcatagttttactgcaattgcatatgagactatgccc 1983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3039 CTCATGAGCTTTGAGGAGTCTTGTATGATTTTCAAGCTGTTGCTTGGTCCATGCGC 3098
OY 1984 alcgaaaaacgctcgttttgcctcgaataltccaaagccaaalcgagggctgcgcat 2043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3099 GTGGGGCAAGTCAGTTTCTGCTGACTGCTGACATGCCAAGCCAAATATACAGAGCCAC 3158
OY 2044 cgtttgctcgttttgaaaaaagaaccaaataagacagccgacgacgaagaaggaaaaag 2103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3159 ATCATCATGATATTTGAAAAAACCCTTTGATTGACAGCTACAGCAGGAGCCCTTAATG 3218
OY 2104 ccaagacacatgtaagggaatttagagtttcgagaagctcttctctatccatgctgc 2163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3219 CCGAACACATTTGGAAGGAATATCATATTTGTTGTAATTCATCATATCCACCCCA 3278
OY 2164 ccaagatgtttatccctccgctggtctatccctcagatltatgacgagaagaacagtagca 2223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3279 CCGGACATCCAGTGTCTTTCAGGAGACTGAGCCTGAGAGTGACAGAGAGCCAGACGCTGCT 3338
OY 2224 ttgtggggaagcagcgctgtgtggaaagacacttctgttcaactctctgagaagacttat 2283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3339 CTGTGAGAGAGAGTGTGCTGTGGGAAGACAGTGTGCTCCTCTGTGAGGGGTTCTTAC 3398
OY 2284 gaccccgtaaggaagaagtgctgtttgtatgtgtgagatgacaaagaatltgaatgacag 2343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3399 GACCCCTTGGGAGGAGAAAGTGTGCTTGTGTCGCAAAAGAAATTAAGCGATGATGTTACG 3458
OY 2344 tggctcgttccccaatagcaatcgttcccaagaagcctgtccttccatgacagact 2403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3459 TGGCTCCGAGACACCTGGGACATCGTGTCCAGAGACCCATCCTGTTGATCGACACATT 3518
OY 2404 gctgagaacatcgtcctatgtgtgacaaagcgtgtgtgtcattagatgagatcaaaaga 2463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3519 GCTGAGAACTTGTGCTATGTGAGAACACCGCGGTGTGTACAGGAAAGATGTGTGAGG 3578
OY 2464 ggcgcgaatgacgaataatcattcatttgaaggtctccctgagaataataacacaga 2523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3579 GCGACAAAGGAGGCCAATACATGCTTTCATCGAGTCACTGCTTAATTAATTAATTAACACT 3638
OY 2524 caagtgtgacgtgaagagacacagcttcttcgagcgcagaaacaaagactaactatgtga 2583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3639 AAGATGAGAGAACAAAGAACTCAGCTCTGTGGGCGAGAAACAAAGCATTTGCGATAGCT 3698
OY 2584 agggctcttctccaaaacccaaaatttattgtgtgagtaggccaactcagacctcgat 2643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3699 CGTGCCCTTGTAGACAGCTCATATATTGCTTTGGATGAAGCCCGTCAAGCTTGAT 3758
OY 2644 aatgacagtgagaaagtgtgtcagcagtcgctgtgataagcagagcaggaagaagaatgc 2703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3759 ACAGAAAGTGAAGAAAGTGTTCAGAGAGCCCTTGAGACAAAGCCAGAGAGAGCCGACCTGCG 3818
OY 2704 ctatgtgactcacaagcgtctctgcaattcagaacgcagattgtatgtgttctgtcag 2763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3819 ATTGTGATTTGCTACCGCTGTGCCACATTCAGAAATGACAGACTTAATAGGGTTTTCAG 3878
OY 2764 aatgagaagaataaagaaagaactatcaagaagctcctgaagaatcagaacataat 2823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3879 AATGGAGAGTCAAGAGAGATGGCAGATGACAGCTGCTGCGACAGAAAGCATCTAT 3938
OY 2824 tttaagttagtgaatgacagctcag 2848

```

```

DB 3939 TTTTCATGTGTCAGTGTCCAGGCTG 3963
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 11
AX012321
LOCUS AX012321 8630 bp DNA PAT 06-SEP-2000
DEFINITION Sequence 2 from Patent EP095374.
ACCESSION AX012321
VERSION AX012321.1 GI:9998370
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 8630)
AUTHORS Baum,C.D., Hildinger,M. and Osterlag,W.P.
TITLE Retroviral vectors for gene transfer
JOURNAL Patent: EP 095374-A 2 10-NOV-1999;
HEINRICH PETTE INST (DE)
FEATURES
location/Qualifiers
source 1..8630
/organism="synthetic construct"
/db_xref="taxon:32630"
/feature="Proximal Plasmid-DNA"
misc_feature 1..160
/feature="Plasmid-R ckgat (puc)"
misc_feature 1..8630
/feature="Retroviral R Expressionsvektor SPbeta9MSAI"
misc_feature 161..677
/feature="5'-LTR"
5'UTR 532..1219
mat_peptide 1220..5062
/feature="MSAI [mdr] cDNA"
misc_feature 5215..5774
/feature="3'-LTR"
misc_feature 5775..8630
/feature="Plasmid-R ckgat (puc)"
BASE COUNT 2234 a 2011 c 2189 g 2196 t
ORIGIN
Query Match 36.0%; Score 1028.6; DB 6; Length 8630;
Best Local Similarity 62.2%; Pred. No. 6.3e-198;
Matches 1694; Conservative 0; Mismatches 989; Indels 42; Gaps 3;
OY 166 attattgaaagaacccagatagataacttccacagctgatatataaactgaatcc 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2321 ATTAATTGATTAATTAAGCCAGATATTGACAGCTATTTCGAAAGAGTGGGCACAAACGATTAAT 2380
OY 226 atagaaggaactgtaaatlttaaaatgttcttcaattatccatcaagaacatctatc 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2381 ATTAAGGGAATTTGGAATTCAGAAATGTTCACTTGTAGTACCATCTCGAANAAGATT 2440
OY 286 aagattctaaagtgatcgaatccgaatlaagctctgagaagaagcgccttgcgtgcgt 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2441 AAGATCTTAAGGGCTGTAACCTGAAGTGAAGTGAAGTGGGAGCGGTGCGCTGATGGA 2500
OY 346 ctcaatgagcagtgaggagaagtagcagtagtccagctctgcagaggttataatgacccgat 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2501 AACAGTGGCTGTGGGAAGGACACACAGTCCAGCTGATGACAGGCTCTATGACCCACA 2560
OY 406 gatgcttataatgtgtgataagaaatgaatcagaatcagaagccttaattgtgcgcatatga 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2561 GAGGGGATGTGCTGTGATGACAGGATTAATGACATTAATTAATTAAGGTTTCTACGG 2620
OY 466 gaccatattgagagtgatgtaagaagcctgtttgttcgggacacacacacacacacacac 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2621 GAATCATTTGTTGTGTGATGACAGAACTGATTTGTTGCCACGATGATGATGATGATGAT 2680
OY 526 atcaaatatgagagagatgagtagtgcattgaaagagatgagagagagcagaaggaagaac 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2681 ATTCCGTATGCGCGTGAATAATGTCACCATGATGATGATGATGATGATGATGATGATGATGAT 2740

```



OY	586	aagcgcgatgattttatcatcaggggttctctataaatttaacttggtaggggaaaaa	645
Db	2741	AATGCCATATACCTTTATCATGAAGAAACGCGCTCATATAATTTTGACACCCCTGGTTGGAGAGGA	2800
OY	646	ggaacccaatagtagtgaaggcgaaagacagatcgcgaattggcgtgaccttaagtga	705
Db	2801	GGGGCCCAAGTTGAGTGGTGGGCGAAGCAGAGAGATCGCATTGCACGtGGCCCTGGTTGCG	2860
OY	706	aacccaagaattcgbatlttagatgaagctagcgtctgccttgatctgaagaagcaagta	765
Db	2861	AAACCCAAAGTCCCTCCTGGATGTAGGCGCACGTGACGCTTGAGACAGAAAGCGAAGCA	2920
OY	766	gctgttcaagctgcacgcgaggaagcgagcaaaagtccggaactcaaatcgtgtgacaaac	825
Db	2921	GTGGTTACAGGTGGCTCTGGATTAAGGCGCAAAAAAGCGGACACCATTTGTATAGCTCAT	2980
OY	826	cgacttctactacticgaagtgcagatttgatgtgacccctaaagagatgagatgctgcg	885
Db	2981	CGTTTGCTTACAGTTGCTCATGTCTGACGTCATCCTCGGTTTGATGATGAGATCATTTGTG	3040
OY	886	gagaaagagacacatgctgtaacctaagcgcaaaaacgagatctatattcaacttgtagt	945
Db	3041	GAGAAAGAAATCATGTATGAACTCATGAAAGAAAGAAAGCAATTTACTTCAAACTTGTCACA	3100
OY	946	tcaagagatatlaaaaaagctgtgaacagatgagtgcaatgaacatctca-----	997
Db	3101	ATGCGACAGCAGCAGCAAAATGTGAATTAAGAAAATGCGACGCTGATGAATCCAAAAAGTGA	3160
OY	998	-----ctgaagaagaagccaactcactcctcctgcactcgtgaaagagcalca---	1045
Db	3161	ATTGATGCCCTTGGAAATGTCTTAAATGATTCAGATCCAGTCTTAATAGAAAAAGATCA	3220
OY	1046	-----agtcagaactcaatggaagcgctgagaaatccacccaactct	1086
Db	3221	ACTCGTAGAGATGCGGTGATACACAAGCCCAAGACAGAAACCTTAGTACCAAGAGCGCT	3280
OY	1087	aaagaagatagctctcctcgtgaagctctctcatataaaatttaagttaacaagcctgaa	1146
Db	3281	CTGGATAAAGTAAATACCTCCAGTTTCTTTTGGAGGATTATTAACCTAAATTTTAACGTAA	3340
OY	1147	tggccttcttggtgtcggggagacattggctcgtgtctaaatggaacgcttcacagta	1206
Db	3341	TGGCCTTATTTTGTGTGGTGTGATTTTGTGCCATTATTAATGGAGGCGCTGCACACACGA	3400
OY	1207	tttccatcatctttgcaaaaaattalaacaagtlttggaataatgataaaacc---aca	1263
Db	3401	TTTGCATATATATTTTCAAGATTATAGGGGTTTTACAAGAATTTGATGATCCTGAAMCA	3460
OY	1264	ttaagaatgatcagaanaattatccatgatatatccgatattcgtcaatttgggtgtatttgctt	1323
Db	3461	AAACGACAGATATAGTAACTTTTTCACATATGTTTCTTGACCCCTTGGAAATTTATTTCTTTT	3520
OY	1324	gtcagttatttcaagaaggagattatttccagcgagaagcgaggaaattttaagatgaga	1383
Db	3521	ATTTCATTTTTCCTTCAAGTTTTCACATTTTGGCAAAAGCTGGGAAATCTCACCAAGCGG	3560
OY	1384	ttaagaaccttggccttcaagaagcactgtatatacagagatattgccttggtttgatgaaag	1443
Db	3581	CTCGATACATAGTGTTTCCGATCCAGTGCACAGACAGAGATGATGGTTGATGATACCCCT	3640
OY	1444	gaaacacgaagaagagccttgacaacaatatagccatagatatagacaanaattcaaga	1503
Db	3701	GCTATAGGTTCCAGGCTTGGCTGTAAATTACCCAAATATGCAAAATCTTGGGACAGAAATA	3760
OY	1564	atcatcttccttaataatgtagatggagagatgacatccctgattctgagatattgctccagta	1623
Db	3761	ATTATATCCTTCACTATATGTTGGCAACTRAACACTTTTACTTTAGCAATTTGTACCCATC	3820
OY	1624	cttgcgtgtgacaggaatgatatgaaccgagacatgactgacgtatttgcacaaagaatag	1683

Db	3621	ATTGCAATACAGAGAGCTGTTGTAATGAAATGTTGTCTGGCAAGCAAGCACTGAAGAATATAG	3860
QY	1684	caagaacttaagaatcctggaaagaatagaacttgaagctcttggagaatatacgtactata	1743
Db	3681	AAAGAACTACGAAGTCTCTGGGAAGATCGCTACTGAAAGCAATGAAAACTTCCGAACGGTT	3940
QY	1744	gtgtcatlaacaaggaaaaagcccttcgagcaaatgtatgaaagaaagtcttcaagctcaa	1803
Db	3941	GTTTCTTTGGACTCAGAGAGCGAGAAATTGAAACATATCTATGCTACAGCTTCCAGGTACCA	4000
QY	1804	caagaataaccctcgaagaagaacagatatttgaagtgtaagtgtaagcttaagcatgccc	1863
Db	4001	TACGGAATCTTTGAGGAAGGACACACATCTTGGAAATTACATTTCTTCCACCCAGCA	4060
QY	1864	tttataatttgcctatgagcaaggttctgatatgtgaagccataatttaattcaagcttga	1923
Db	4061	ATGATGATATTTTTCATACGTGTGATGATTTTCCGGTTTGGAGCCTACTTGGTGGACATATA	4120
QY	1924	cgaaatgaccccaaggagcgtcatagitttctactgcgaattgtcatatgtagagctatggcc	1983
Db	4121	CTCATGAGCTTTTGAAGATGTTCTGTTAGTATTTTACAGCTGTTGTCTTGGTGCATGGCC	4180
QY	1984	atcggaaaaacgctcgttcttggtcctctgataatccaaagccaatccggggcgctggcat	2043
Db	4181	GTGGGGCAACTCAGTTCATTTGCTCTGTGACTATGCCAAAGCCAAATATACAGCAGCCAC	4240
QY	2044	ctgttgcctctgttggaaaagaaccaaatagagaacgagccgagtcgaagaaggaaag	2103
Db	4241	ATCATCATGATCATTTGAAAAAACCCCTTTGATTTGACAGCTACAGCAGCGAAGCCCTAATG	4300
QY	2104	ccagacacatgtgaaggaaattagaagtttcgagaagtcctcttctatccatctgcgc	2163
Db	4301	CCGAAACCATTTGGAAGGAATGTCACATTTTGGTGAAGTTGTATTCATCATCCACCCGA	4360
QY	2164	ccgaatgtttcaactcctcgttgctctataccctcagatattgagcgaggaaagaagtagca	2223
Db	4361	CCGGAACATCCAGTCTTCAGGGACTGAGCCTGGAAGTAAAGAGGCGACAGCGCTGCT	4420
QY	2224	tttttgaggagacagcgcgctcttggaaaagacacttctgtcaactctctgcagaagacttat	2283
Db	4421	CTGTGTGGCAGCAGTGTGCTGTGGGAAGACACAGTGTGCACACTCTTGAGGGGTTTAC	4480
QY	2284	gaecccgtgcagaagcaaatgtcgtcttgaatgtgtatgtatgaataaagaatigaatgtacag	2343
Db	4481	GACCCCTTTGGCAGGAAAGTGTCTTGTGATGGCAAGAAATTAACCGACTGATATTTTCAG	4540
QY	2344	tggtctcggttcccaaatagaatcctcttccaaagagccgtgtccttctaactcgaagcatt	2403
Db	4541	TGCGTCCGACACACCCTGGGCATCTGTGTCTCCAGAGACCCTATCTCTTTGACTGCGACATY	4600
QY	2404	gcttaagaacatcgccctatgttgaacaacagcgctgtgtgtccattagatgaatcaaaaga	2463
Db	4601	GCTTAGAACATTTGCCATATGAGACACACACCGGGTGTGTCTCACAGGAAGATGTGTAGG	4660
QY	2464	gcgcgaatgtcagcaaatatccattcttattatgaaggtctccctgagaaatacaaca	2523
Db	4661	GCACGAAGAGAGCGCAACATACATATGCTTATCTCATCTGACTCTCTTAATAATATAGCACT	4720
QY	2524	caagtgtactgaagaaggagcacagcttcttcgcgccgagaatacaagctatgtctattgca	2583
Db	4721	AAATAGGAACAAAGAAATCTGACGCTCTGTGGGCGCAAAACAGGATTTGCCATTAACCT	4780
QY	2584	aggcgtcttccaaaacccaaatattatgttgaatgaagcgaacttcgaagccctgat	2643
Db	4781	CGTCCCTTGTATGACAGCCTCATATTTTGTCTTTTGGATGAAGCCACAGCTCAAGCTTGAT	4840
QY	2644	aatcacagttagaagatgtgttcagcatcgcccttgataaagccaagcagaagcggagaagcatgc	2703
Db	4841	ACGAAGAAGTAAAAAGTTGTTCACGAAGACGCTGTGAGCAAGAACCGCGCACCTGAC	4900
QY	2704	ctagtgtcaactcaagcgtctctctgcaattcagaacgcgaattgtatgtgtgtctgcac	2763



QY	1753	acaaagaaagacgttcagcaaatggttgaagagatgcttcagatccaaacagaat	1813
Db	2760	ACCCAGGAAGAAATTTGATCAATGTAATTGTTGAAAAATTTATGACCTTACAGAAAT	2819
QY	1813	acctcgaaagacacagatatttgaagctgttatgcattcaagccatgacctatata	1872
Db	2820	TCTGTGACGAAGGACACACATCTATGGAATTACTTTAGATCTCACAGCACTTATGAT	2879
QY	1873	tttgccattcagcaggttttgatttgaagcttaattcaagctcgagcaatgacc	1932
Db	2880	TTTTCTATCCCGGTGTGTTTCATTTGGTATCTCATTTGTGAATGACATATGGCG	2939
QY	1933	ccaaaggagcatgctacatgatttttactgaattgtacatgtagcctatgccaatga	1992
Db	2940	TTTCAGACATGTTATTTTCGGTGTTTCTGCAATTTGTAATTTGGTGCAGTGGCTTAGACAT	2999
QY	1993	acgctcgltttgctcctgaataatccaaagccaatcgagggtcgcatctgtttcc	2052
Db	3000	GCCAGTTCATTTCTCCAGACTATGCTAAAGCTGTCAGCCACCTTATTCATG	3059
QY	2053	tttgttgaagaagaaccacaatatagacagccgcatccaagaaggagaaagccaagca	2112
Db	3060	CTGTTTAAAGACAAACCTCTGATTGACACTACAGTGAAGAGGGCTGAAGCTTGATAA	3119
QY	2113	tgtgaaggaaattagaatttcagaagaagctctcttctatccatgctgcgccagatgt	2172
Db	3120	TTTTAAAGAAATTTAACTATTATTAAGATCGTGTTCCATCTATCCACCGGACAAACGTG	3179
QY	2173	ttaactcctcggtgcattatcccccagtatcttgagcgagaagaacagtgatgttggg	2232
Db	3180	CCAAGTCTTAGAGGGCTGAAGCTCGAGGCTGTAAGAAAGGCCAACAACCTATGCGTGTGGCG	3239
QY	2233	agccagcgctgttggagaagaacactctctgtcactctctgcaagagacttatgaaacccgt	2292
Db	3240	AGCAGTGGCTGTGGGAAGACACAGGTGATCCAGCTCGAGCGGTGTACAGACCCCTTG	3299
QY	2293	caaggacaagtgcctgttttgaatgtgtgatgcaaaagaatbgaatgtaagtgctcgt	2352
Db	3300	GGGGGACAGTGTCTTCGATGTGTCAAGAAAGCAAAACAACTAAATGTCAGTGGCTCGA	3359
QY	2353	tcccaaatagcaatcgcttcccaagaagcgttgctcttaactgaagatbgaatgtgaagac	2412
Db	3360	GCTAACTCGGAATCGTGTCAGAGGACCTATTCATTGATCGACAGCTTCCAGAAAT	3419
QY	2413	atcgctatgttgacaaacagccgtgtgtgtgcaattagatlgatcaaaagaagccgcaat	2472
Db	3420	ATTCTCATGTGAGACAAACACCCGGTTGTATCACAGAGTAATGTGATGCACCCAA	3479
QY	2473	gcagcaaatlccattcttatttgaagttlccctgagaaatatacaacaaagtgtga	2532
Db	3480	GCTCCCAACATACCTCTTTCATCTGAGAGCTTACCCCAATAATATGAAACAAGAGTGGGA	3539
QY	2533	ctgaagaagacacagcttcttgccgcgcgaagaacaagaactagctatttgaagggctct	2592
Db	3540	GATTAAGGAGACTGAGCTCTCAGAGAGGTCAAAAAACAGAGATTGTATTTGCCGAGCCCTC	3599
QY	2593	ctccaaaaccacaatttatttgttggatgagggcaacttcagccctcgatatagaagt	2652
Db	3600	ATTCAGAACCTCAAAATCCCTCGTTGGATGAAGCTACATCAAGCTCTGGATTCCTAAAGT	3659
QY	2653	gagaagtggttcagcatgcttctgtataagaccaggaagcggaagacatgcttaagtc	2712
Db	3660	GAATAGGTGTCCAAAGAACGCTTGACAAAGCCAGAGAAAGCCCGACCTCGCATTTGTGATT	3719
QY	2713	actcaagcgtctctgtcaatttcgaacgcagaatttgaatggtttctgcacatggaag	2772
Db	3720	GCTCACCGCTGTCACACATTCAGAAATGACAGATTAAATAGTGGTTCAGAAATGGGAGA	3779
QY	2773	ataaagaacaaagaaactctataagagcttccgtagaagaatcagagcatatatttaagta	2833
Db	3780	GTCAAGAGCATGTGCACGATCGATCGACAGCTGTCTGGCAACAAGAGGACATATTTTTCATGT	3839

RESULT	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM
3y	2833	gtgaatgcacagtcagttc	2850			
Db	3840	gtcagcttcacagcctggc	3857			
RESULT	13					
LOCUS	HUMMDR3	3924 bp	mRNA	PRI	11-JUN-1993	
DEFINITION	Human membrane glycoprotein P (mdr3) mRNA, complete cds.					
ACCESSION	M23234					
KEYWORDS	M23234.1 GI:187501					
SOURCE	P-glycoprotein; membrane glycoprotein.					
ORGANISM	Human liver, cDNA to mRNA, clone 3.27.					
REFERENCE	1 (bases 1 to 3924)					
AUTHORS	van der Bliek, A.M., Koolman, P.M., Schneider, C., and Borst, P. A.					
JOURNAL	Sequence of mdr3 cDNA encoding a human P-glycoprotein					
MEDLINE	89138016					
COMMENT	Draft entry and computer-readable sequence for [1] kindly provided by P. Borst, 21-MAR-1989.					
FEATURES						
source	location=Qualifiers					
mrna	1..3924					
cds	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	<1..3824					
	/product="mdr3 mRNA"					
	33..3872					
	/note="P-glycoprotein"					
	/codon_start=1					
	/protein_id="AAA36207.1"					
	/db_xref="GI:307181"					
	translation=MDLEPAKNGTAMPTNASEGDFELGJSKQRRKTKTYKMGVLT					
	LEFRSDWDKLFMSLGTIMAIAHSGSLPLMAIVFGEWTDKFDVADGNSEPFVNSISL					
	LNPGRILIEEENETRYAAYYSIGAGVLAAYIAYSFMTIAGROIRKIRKPFNALIRQ					
	ENGFEDINDNTTELNTLRLLDPSIGISEGKGVGMFOVAATFEGATVGFIRGKLTIL					
	VYMAISPIILISAAMVAKIISAFDSKELAAKAGAAVEELGAPRTVIAFGGONKEL					
	ERYKRIHENKEIKIKKAKISANISMGATALLIYASVALARWGSYTIYSKYETGNM					
	TVFSPILIGASVGOAAPCDIDAFANAGAAVIFDIIIDNNPKIDSESEKHPDPSIG					
	NLEINDVAFSPYSRANVKILKGLMKVQSGQVALVSSGCKSTYQILQRLYPDE					
	GTILINDQDIRENFVNNTLRLEITIGVSOEPLYFTTILAEINICYRGANTMDIKAVE					
	ANAEYFELPKQKEDTLVGERGAQLSGGOKRIIAEALVRNPFLILIDDEATSLDTE					
	SEAVOALAIKAPRGRTTIVIAHRLSTVRNADYIAGDEQVIVQSGHSELMKKEGY					
	EKLVMQTSQSQIDSEFELNDEKAARMAPNMGKSLPASHSDKNIKNSOMOKSID					
	VEDIDLEANPVSEFLVKLTKEMTFYVAGTATANGLOPABSVITSEPIALRG					
	PGDVAQKQCNIFSLIFLGIIISFTPLQGFTEPKAGETILTRLRSAFAMLRQ					
	MDLFWDDHKNTGALSTRLATDAQVQAGATGTRALIAQNIAMLGTLIIISFYQWL					
	TLLAVPIITIAVGIYEMKILAGNARKDKKELPAEAKIATEALINIRTVASLQOEK					
	FESYVVEIKGVANVSOKAHVITFTPSISOAEYFAGCFREGAVILVNGHFRPDD					
	VIIYFSAIVGCAVLAHGAASSFPAPYAKAKISAAILFMLFERPRLIDSEREGKPKF					
	EGNITREHVVFNNTYTRANVPYLOGSLSEVKKQTLALVSSGCKSTVYOLLREFPD					
	LAGVILLDGEAKRLNVOQLRAQLGIVSOEPLTIDCSIAENIAIGDNRVSDSEIVS					
	AAKNAITHPEITELPHKYEYRVGDKGTQLSGGOKRIIAEALIRQPOLILLDEATSA					
	LIDSEKVEQALDKAREGRTCIYIAHRLSTIQNADLVIVFONGRYVEHETHOOLLAO					
	KGIYFSWVSQAGRONL"					
BASE COUNT	1145 a	790 c	977 g	1012 t		
ORIGIN						
Query Match	36.0%; Score 1027.2; DB 9; Length 3924;					
Best Local Similarity	61.9%; Pred. No. 1.2e-197;					
Matches 1682; Conservative	0; Mismatches 1003; Indels 33; Gaps 2;					
3y	166	attatgataagaaccacagtatagataactttccacacgcttgatatataaacctgatacc	225			
Db	1140	ATTATTGATTAATAAATCCATAAAATTGACAGTTTTCAGAGAGACGACACAAACCAACAGC	1199			
3y	226	atagaagaacgtgtaagatttaaaatttcttcattcatatccataagaaccatcatc	285			
Db	1200	ATCAAAGGAATTTGGAGTTCATGATCATGCTTCTTCTTACCCCTTCGAGACTTAAGCTC	1259			

Oy	286	aaggtctggaagagtcgtaatctcagaalttaagtctcgsgaagaaagtcgcttgctg	345
Db	1280	AAGATCTTTGAAGGGCCCTCAACCTGTAAGGTGTCAGAGTGGGACACGGTGGCCCTGGTTGGA	1319
Oy	346	ctcaatgicagctgggaagaaagtcagtgtaagtcgaactcttcagagtgtaatgatccgcat	405
Db	1320	AGTAGTGCGCTGTGGGAAGACACAAAGGTCGACGCTATACAGAGGCTCTATACCTTGAT	1379
Oy	406	gatygtcttaacabgtytgaatgaagatcagaacttaaatlytgygcataltcga	465
Db	1380	GAGGGCAACAAATTAACTGATGGGGCAGAGATATAGAACCTTTAATGTAAACTATCTGAGG	1439
Oy	466	gaccatattgagatggttgaatcaagagcctglttlytgcggaaaccaatcaagtaacat	525
Db	1440	GAATTCATTTGGTGGTGGAGATCTCAGAGCGCGTCTCTTTTCCACCACTTCCTCAAAAT	1499
Oy	526	atcaagatcagcagagatgtgtgactcgaatgaagagatgagagagcaagaggaagca	585
Db	1500	ATTGTGTTATGGCCGTGGAAATGTAAACCATGTGATGAATTAAGAAAGCTGTCAAGAGGCC	1555
Oy	586	aatgcgatatttaatacatcagtagtcttccataaatttaatacatctgylaagggaaaaa	645
Db	1560	AACCGCTATGAGTTTATCATGAAATTAACCAAGAAATTTTGACCCCTGGTTGAGAGAGA	1619
Oy	646	ggaagctcaaatgagtygagggcagaagaacagaggaatcgcaattgctgcgtccttaagtcga	705
Db	1620	GGGGCCACAGCTGAGTGTGTGGGCAAGACAAAGGATGCGCATTCGACAGTGCCTCGTGGTGGC	1679
Oy	706	aaccacaagattcgaattttagatgagagctgaagtcgcgccttgattcagaagaagca	765
Db	1680	AACCCCAAGATCCTTCTGCTGGATGAGGCCACGCTCAGCATTTGGACACAGAAAGTGAAGCT	1733
Oy	766	gctgttcaagctgcagctcaggaagagcgagcaaaagctcgcgactacaatcgttgglaagcaac	825
Db	1740	GAGGTACAGGACGCTGTGATTAAGGCCAGAGAAAGGCCGCGGACCACTTGATGATAGCACAC	1799
Oy	826	cgacttctactatctcgaagtgcgaattgattgtgacctaaaggaatgaaatgctgcg	885
Db	1800	CGAGTGTCTTAAGGGTCCCAAAATGCAAGATGTCATGCTGGGTTTGAGAGATGAGCTAATTG	1855
Oy	886	gagaagaagagacatgctgtaacatcagcgaacagaggtcctatattcacttactgta-	944
Db	1860	GAGCAGAGGAAGCCACACGCGAATGATGAAAGAAAGGGGCTGTACTTCAAACTGTGTACAC	1919
Oy	945	-----gtcacagatattaanaaagctatataacagatg	978
Db	1920	ATGCAAGACATCAGGAACCCAGATCCAGTCCAGTAAGAAATTTGAACTAAATGTGAAAAAGCT	1979
Oy	979	gagtcaatgacatatcttactgaagaagaacccaactaactctcctcgtcactcgtlgaag	1038
Db	1980	GCCACTGTGAATGGCCCCAAATGCGTGGAAATCTCGGCTTTTAGGCACTTCTACCTAGAAA	2033
Oy	1039	agcatcaagtcgaactctcaltgacaagc-----tgaagatccacccaactctaagag	1092
Db	2040	AACCTTAAAAATTCACAAATGTCTCAGAAAGCGCTTGATGTGGAAACCCATGAGACTTGAA	2099
Oy	1093	ataagcttcctcgaagctcctctataaanaatttaagaattlaaacaagcgttaatgacct	1152
Db	2100	GCAATATGGCCACGACGAGTCTCTTTCGAAGGTCTTAACCTGAATTAACCAAGATGGGCC	2155
Oy	1153	cttgtygtctcgtgggacatggtcctcgttctaataatgtaactgttcaatcagtaatttcc	1212
Db	2160	TACTTTGTGCGTGGGGAACAGATGTGCCATGTCCAAAGGGGGGCTTCAGCGCGCATTTTCA	2219
Oy	1213	atcgtcttgcaaaaattataaccatgitttggaaataatgataataaaccacataaagcat	1274
Db	2220	GTCAATATTCAGAGATCAACGCAATTTTGGACCAAGCGCATGATGCACTGAAGCAAGAG	2279
Oy	1273	gatcagaagaatttatccacatgataatcgtcaattcttggtgttaattgcttgcagttat	1332
Db	2280	AAGTGCACAAATATCTCTTTGATTTTCTTATTTCTGTGGGAATATTTCTTTTTCATTTC	2333
Oy	1333	ttcatcgaaggatatttlaacgcagagcaagggaaatttlaacgaatgaaatlaagac	1392

Db	2340	TTCCCTTCAAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTTACCAAGAAGACTGGCGTCA	2399
Oy	1393	ttggccttcaaaagcagctatatactcaggaatattgcctgtgttgatgaaaggaanaacagc	1452
Db	2400	ATGGCTTTTAAAGCAATGCTAAGACAGAGACATGAGCTGGTGTGTGATGACCATTAATAACAGT	2459
Oy	1453	acaggaagccttgacacacatatagccatagatatagcaaaatccaaggaacagagt	1512
Db	2460	ACTGGTGCACCTTCTTACAMACCTTGGCCAGAGATGCGCCCAAGTCCAAAGGACCACAGA	2519
Oy	1513	tccaggaattgagcgttaacacaaatgcacctaacaatgaggaacttcaagtaattcc	1572
Db	2520	ACCGAGTTGGCTTTAATTTGACAGAAATPAGCTTACCTTTGGAACGTGTATTATCAATATCA	2579
Oy	1573	tttatatagatggagagatgcacatccgtatctcgtatctgaattgcctcagtaactgcg	1632
Db	2580	TTTATCTACAGGTTTGCGAGTTTAAACCCATATGCTATTAGCAGTGTGTTCCATTATTCGTGTG	2639
Oy	1633	acaggaatlgatlgaaacccgcagcaatgactggaattgcacaagaataagcaagaact	1692
Db	2640	TCAGGAATTTGTTGAATAATGAAATTTGGTGGCCGAAATGCCAAAGAAGATPAAAATAAGACTG	2699
Oy	1693	aagcattctgtaaaagataagaactggaagctttggagaataatagctactatagtgcatata	1752
Db	2700	GAACTGCTGGAAAGATTTGGACAGAGGCAATPAGAAAATTTAAGACAGTGTGTCTTTTG	2759
Oy	1753	acaaggaanaaagccttcgcagcaaatgtaatgaaagatgcttcagacccaacacagaat	1812
Db	2760	ACCCAGGAATAAATTTGAAATCAATATGTTGTAATAATTTGTATGAGACTTACAGGAAAT	2819
Oy	1813	acctogaanaagcagatlatltggaagctgltatgcatccaagcaatgcttatatat	1872
Db	2820	TCTGTGCAGAGGACACACATCTPAGGAATPACTTTAGATCTCACAAGCATTTATGTAT	2879
Oy	1873	tttgcctatgagcaaggtttgcgatttgagcctatattatccaagcttggagaatggcc	1932
Db	2880	TTTTCTPAGGCGGTTTTCATTTGTCATTTGGTCATATCTAATGTGAATGAGATATTTGGCG	2939
Oy	1933	ccaagag9cacgttcacagtttctactgcaatgcatagagcattagagcattgccaatgaa	1992
Db	2940	TTTCAGAGATGTTATTTCTGGTGTTTCTGCAATTGTATTTGGTGACAGTGGCTGTAGGAAAT	2999
Oy	1993	acgctcgttttgycctcctgataatccaagaagcaaaatcg9gg9cgtgycatcgttttgc	2052
Db	3000	GCCAGTTTCATTTGCTCCAGCATPAGCTAAGCTGTACAGCTGTACAGCCACTTATTCATG	3059
Oy	2053	ttgttgaaanaagaaaccaaatatagacagccgagtcgaaggaaggaanaaagccagaca	2112
Db	3060	CTGTTTGAATAACACCTCTGATTTGACGTACAGTGAAGAGAGGGGCTGMAAGCTCATATAA	3119
Oy	2113	tgtbaaggaatlttagagtttgaggaagcctcttctctatccaatgctgcgcagatgt	2172
Db	3120	TTTAAAGAAATATPACATTTAATPAGAGTGTGTTTAACATPACCACCCGAGCAAAAGTGTG	3179
Oy	2173	ttcatccctcgttgctatccctccagfatlgac9gagaaagacagtagtactltg9g	2232
Db	3180	CCAATGCTTCAGGGGCTGAGCCTGGAGGTGAATAAAGCCAGACACTAGCCCTGTGGTGGC	3239
Oy	2233	agcagcgctgtg9gaaagcactctgttcaactctgcagagacattatgaaccgctg	2292
Db	3240	AGCAGTGGCTGTGGAGAGACACAGGTGTGTCCACTCTTGAGCGGGTTCACGACCCCTTGTG	3299
Oy	2293	caagagcaaaagtcgttggltggtgltg9gaaagaaagtltgaatgacagtgagtcgt	2352
Db	3300	GGGGGACAGGCTTCTCGATGTGCAAGAGCAAGCAAAATACTAATGTCCAGTGGCTTCAGA	3359
Oy	2353	ttcccaaatagcaatcgttcttccaagagcctgtgtcttcaactgtgacgatltgtgaagac	2412
Db	3360	GCTCAACGTGGGAATCGTGTCTCAGAGAGCCTATCTTTGACTGGCAGCATTTGCCGAGAT	3419
Oy	2413	atcgactatgltgacaacagcgtgtgtgtgccaattagatgagatcaagaagccgacaat	2472

Db 3420 ATTGCTATGAGACAAACAGCGGGCTGTATACAGAGATGAATTTGTGATGACAGCCAAA 3479  
Qy 2473 gcaacaatcatcattcttattgaagctccctgagaataacaacacaagttbga 2532  
Db 3480 GCTCCACATACATACATCTTTTATGAGACGTTACCCCAATATGAAACAAAGAGTGGGA 3539  
Qy 2533 ctgaagagacacagcttctgagcgccagaacaaagactagctatttgaagagctctt 2592  
Db 3540 GATGAAGGGACTCAGCTCTCAGAGAGTCATAAACAAGAGATGTGCTATTTGCCGAGCCCTC 3599  
Qy 2593 ctccaaaacccaataattttatttatttgaatgaagcacttaagccctclogataatgacagt 2652  
Db 3600 ATCGACAGACCTCAAACTCTCTGTGATGAACCTACATCACTCTGTGATGATGAAGT 3659  
Qy 2653 gagaagtggttcacagctcccttgataaagccagagacggaagagacatgctagtgtc 2712  
Db 3660 GAAAGAGTGTCCAAAGAAAGCCCTGGACAAAGCCAGAGAGAGCCGACCTGCATTTGTGATT 3719  
Qy 2713 actcagaagctctctgaatctcaagaacgacagatttgaatgatttctgacaaatggaag 2772  
Db 3720 GCTCAGCGCTGTCCACCATCCAGATGACACTTAATAGTGTGTTTCAGAAATGGGAGA 3779  
Qy 2773 ataaaggaacaaagaactcaatcaagagctcctgagaatcgaacataatatttaagtta 2832  
Db 3780 GTCAGAGGCAATGGCAGCAGCAGCAGCTGCTGGCACAGAAAGCAATCTATTTTCAATG 3839  
Qy 2833 gtgaatgcacagctgagt 2850  
Db 3840 GTCAGTGTCCAGGCTGGG 3857

RESULT 14  
AX012320 LOCUS AX012320 8630 bp DNA  
DEFINITION Sequence 1 from Patent EP0955374. PAT 06-SEP-2000  
ACCESSION AX012320  
VERSION AX012320.1 GI:9998369  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 8630)  
ARTHOVS Baum,C.D., Hildinger,M. and Osterlag,W.P.  
TITLE Retroviral vectors for gene transfer  
JOURNAL Patent: EP 0955374-A 1 10-NOV-1999;  
HEINRICH PETTE INST (DE)  
FEATURES  
source  
1. .8630  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Proviral Plasmid-DNA"  
misc\_feature 1. .8630  
/note="Retroviral Expressionsvektor Sfbeta7im4"  
misc\_feature 1. .160  
/note="Plasmid-R ckyrat (pUC)"  
misc\_feature 161. .677  
/note="5'-LTR"  
5'UTR 532. .1219  
mat\_peptide 1220. .5062  
misc\_feature 5215. .5774  
/note="3'-LTR"  
misc\_feature 5775. .8630  
/note="Plasmid-R ckyrat (pUC)"  
BASE COUNT 2235 a 2011 c 2187 g 2197 t  
ORIGIN

Query Match 36.0%; Score 1027; DB 6; Length 8630;  
Best Local Similarity 62.1%; Pred. No. 1.3e-197;  
Matches 1693; Conservative 0; Mismatches 990; Indels 42; Gaps 3;  
Qy 166 atattgataagaaaccagatagataactttccacagctgataataaactgaatcc 225

Db 2321 ATAAATTGATTAATTAAGCCCAAGATTAAGACAGCTATTGCAAGAGTGGGCAAAACAGATAAT 2380  
Qy 226 atagaagaaacttgggaatttaaaaatglttcttcaattatcatcaagaacctatc 285  
Db 2381 ATTAAGGCAAAATTTGGAATTCAGAAATGTTTCATCTGATTACCATCTCGAAAGAAATG 2440  
Qy 286 aagatctgaagaggttcgaatctcagaatlaagtcgagagacagtcgcttgcgt 345  
Db 2441 AAGATCTTGAAGGGCCCTGAACCTGAAGTGCAGAGTGGGCGACAGAGTGGCCCTGTTGA 2500  
Qy 346 ctcaatgagcagtggaagagtagcgttagtccagctcttcgaaggttatatgacggat 405  
Db 2501 AACAGTGGCTGTGGGAAGAGACACAACAGTCCAGCTGATGACAGAGCTTATGACCCACA 2560  
Qy 406 gatgcttatacaatggtggaatgagaatgacatcgaagcttaaatgctgacatatacga 465  
Db 2561 GAGGGATGTCAGTGTGATGAGACAGATATTTAGGACCATTAATTAAGCTTTCTACGG 2620  
Qy 466 gaccatattgagtggttagtcaagagcctgttctgctcgaccacatcagtaacaat 525  
Db 2621 GAATTCATTGGTGTGGTGAAGTCAAGAACCTTATTGTTGCCACACGATAGCTGAANAAC 2680  
Qy 526 atcaagatgacgagatgactgactgaatgaagatgagagagagagagagagagagagag 585  
Db 2681 ATTGCTATGCGCGTGAAATATGTCACCATGATGATGATGAAAGAGTGTCAAGAGGCC 2740  
Qy 586 aatgcgtatgatttatacaatgaggttccataataatttaacatggttaggggaaaaa 645  
Db 2741 AATGCTATGACTTTATCATTAACCTGATTAATTTGACACCCCTGTGGAGAGAGA 2800  
Qy 646 ggaagctcaaatgagtgagagcagaagaagagatgcgaattgctgctgctgacttgcga 705  
Db 2801 GGGGCGCCAGTTGATGATGGGCGACAGACAGAGATGCCATTCACAGTGGCTGTTCCG 2860  
Qy 706 aaccccaagattctgattttagatgaggtactgctgctccctgagatcaagaagaatga 765  
Db 2861 AACCCCAAGATCTCTGCTGATGAGGCGACGTGACGCTTGACACAGAAAGCGAAGCA 2920  
Qy 766 gctgttcaagctgcagtggaagagcagagcgaagaagtgcgactaataatcgtgtgacac 825  
Db 2921 GTGGTTACAGTGGCTCTGATTAAGGCCAGAAAGCTGGACACCATTTGTATGATCTAT 2980  
Qy 826 gcaattctactatcgaagtgacagatttgaatgcacaaagagtgagatgctgagc 885  
Db 2981 CGTTTGTCTACAGTTCCGAATGCTGACGTCAATCCCTGTTTCATGATGAGTATTTGTG 3040  
Qy 886 gagaagagacacatgctgaactaagtcgaacaaacgaggtctataatctcaatgagtg 945  
Db 3041 GAGAAAGGAATTCATGATGACATCATGAAAGAGAGCATTTACTTCAAACTGTGACA 3100  
Qy 946 tcaagatattaaaaaagctgataagacagatgagatgaatataatctca----- 997  
Db 3101 ATGACAGACAGCAGGAATGAAGTTAGAAAATCCAGCTGATGATCCAAAAGTGAA 3160  
Qy 998 -----ctgaagaagaagcacaactcaactctctcgaactcgtgaagagacata--- 1045  
Db 3161 ATTGATGCTTGGAAATGTCTTCAAAATGATTCAGATTCACATCAATTAATTAAGAAAGTCA 3220  
Qy 1046 -----agtcagacttcatgacaaagcgtgagaatccaccacatct 1086  
Db 3221 ACTGTAAGAGTGTCCGATGATCACAAGCCCAAGCAAGAAAGTCTAGTACCAAGAGGCT 3280  
Qy 1087 aaagagataagctctctctgaagctctctataaaaaatttaaaagttaaaagacccgtgaa 1146  
Db 3281 CTGGATGAAGATTAATCCCTCCAGTTCTTTGGAGATTAAGAGCTTAATTAATTAAGTAA 3340  
Qy 1147 tggcctttgtggttcggagacatlggctctgttctaaatggaactgttcaacagta 1206  
Db 3341 TGGCTTATTTTGT 3400  
Qy 1207 ttctccatcacttgcgaataataataacatgcttgggaataatgataaac---aca 1263

Db 3401 TTTGCAATATATTTTCAAGATATAGGGGTTTTCACAGAAATGATGATCCTGAACA 3460  
Oy 1264 ttaagcgtatgcagaattatccatgatattcgcatttgggttatttgcctt 1323  
Db 3461 AAGCAGCAAGATAGTACTGTTTTCACATATGTTTTCAGCCCTTGGAATATTTCTTTT 3520  
Oy 1324 gtcaattatccatgcaggaattatccagcagaaggggaattttaaagtgaga 1383  
Db 3521 ATTACATTTTTCCTTCAAGGTTTTCACATTTGGCAAAAGCTGGAGAAATCTCCACAGCG 3580  
Oy 1384 ttaagacacgttgccttcaaaagcattatatacagatattgcctggtttgatgaag 1443  
Db 3581 CTCCATATCATGTTTTCGATCCATGCTCAGACAGATGATGATGGTTGATGACCT 3640  
Oy 1444 gaaacagacagaagggcttgacaacaatatagcattagatagcacaaattcaagga 1503  
Db 3641 AAAAACACCACTGAGCAATTTGATCCAGGCTGCCAATGATGCTCAACTTAAAGG 3700  
Oy 1504 gcaacaggttccaggaattgcgtcttaacaacaatgcaactaagattggaattcagtt 1563  
Db 3701 GCTATAGGTTCCAGGCTTGCCTATATACCCAGATATAGCAAACTTGGGACAGAAATA 3760  
Oy 1564 atcaattccttataatagatggagatagacattcctgattctgagtattgctcagta 1623  
Db 3761 ATTATATCTTCATCTATAGTTGGCACTACTGCTACTCTAGCAATTTGACCCATC 3820  
Oy 1624 ctggcgcgacaggaattggaacccgacgaactgactggttggcacaacaataag 1683  
Db 3821 ATTGCAATAGCGAGGAGTTGTAATAATGTTGCTGCGACAGCACTGAAATATAG 3880  
Oy 1684 caagaaacttaagcatgctgcgagaagaatagaactgaagcttggagaataatgctata 1743  
Db 3881 AAGAGAACTAGAAAGTGTGGGAGAGTCCGCTACTGAAAGCAATGAAACTCTCGAACCTT 3940  
Oy 1744 gtgtcaataacaagggaagaagccttcgagaacaatglatgaagagatgcttcaactaa 1803  
Db 3941 GTTCTTTGACTCAGAGCAGACAGATTGTAACATATGATGCTCAGAGTTTGACATACCA 4000  
Oy 1804 caccgaataactcgaagaagaacacagattatggaactgttatagcattcagcattgcc 1863  
Db 4001 TACAGAAACTCTTTGAGAAAGCACACATCTTTGGAATTTACATTTTCCCTCACCCAGGA 4060  
Oy 1864 ttatataatttgcatagcagcagggattgcagcttggagccattatcaactga 1923  
Db 4061 ATGATGATTTTCTCTATAGCTGAGATGTTCCGGTTTGGAGCCTACTGTTGGCAATAAA 4120  
Oy 1924 cgaatgacccagaggagcattagatttactgcaattgcatatgagactatggcc 1983  
Db 4121 CTCATGACCTTTGAGAGATGTTCTGTATATTTTCAGCTGTGCTTTGGTGCATGGCC 4180  
Oy 1984 atcggaataaagcctgcttgcctcgaataattccaagccaatcggggctgcgcat 2043  
Db 4181 GTGGGGCAGTCACTTATTTGCTCTGACTATGTCGAAGGCCAAATAATACACACCCAC 4240  
Oy 2044 ctgttgccttgcgtgagaagaacaaatataagacacccgagtcagaaggaagaag 2103  
Db 4241 ATTCATCATGATCATTTGAAAAAACCCTTTGATGACAGCTACAGCAGGAAGCCTTAATG 4300  
Oy 2104 ccagacacatgtgaagggaatttagagtttcgagaagtccttcttctatccatgctgc 2163  
Db 4301 CCGAACAATTTGGAAGAAATGTCATATTTGGTGAATTTGATTCACATATCCACCCGA 4360  
Oy 2164 ccaagtgttccatcctccgtggtctatccctcagatattgagcgaaggaagcagtagca 2223  
Db 4361 CCGGACATCCCAAGTCTTCAGGAGTGAAGCCGAGGAGGAAGGCCAGACGCTGGCT 4420  
Oy 2224 ttgtgggagcaagcggctgtgggaaagaacactctgttcaactctcagaagactttat 2283  
Db 4421 CTGGTGGGACAGAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCGAGCGGTTCTAC 4480  
Oy 2284 gaccctgtgcaaggaacaagtctgttgaatggtgtgagatgcagaagaattgaatgtacag 2343  
Db 4481 GACCCCTTGGCAGGGAAGTGTCTGCTGATGCAAAAGAAATTAAGCACTGAATGTTCAAG 4540

Oy 2344 tggctcgttcccaatagcaatgcttctcagaagccctgtccttcaactgcagcatt 2403  
Db 4541 TGGCTCCAGAGCACCTGGGACATGCTGTCCAGGAGCCCATCTCTTTGACTGCGACAT 4600  
Oy 2404 gctggaacatgacctatggtacacacagccgtgtgttcctatagatagatcaaga 2463  
Db 4601 GGTGAGAACATTTGCCATAGAGACAAACAGCCGGGTGTCTCAGAGAAAGATCCTGAG 4660  
Oy 2464 gccgcaatgcaagaatatccattcttatttgaagtgctcctcctgagaatacaaca 2523  
Db 4661 GCAGCAAGAGAGCCCAACATACATCTCTTCATCGATGCTACTCTTAATTAATATGACT 4720  
Oy 2524 caagtgtgacgaaggaagcagacttctgcgcgcgaagaagaagcagctagcttga 2583  
Db 4721 AAAGTGGAGACAAAGAACTGCTGTGTGGTGCACAGAAACAGCACTTCCATAGCT 4780  
Oy 2584 aggcctcttcccaaaacccaatttattgttggatgagagccacttcaagccctgat 2643  
Db 4781 CGTGCCCTTGTAGACAGCTCATATTTTGTCTTTGGATGAAAGCCACGCTCCTGGAT 4840  
Oy 2644 aatgacagtgagaaggtgtgtcaagcatgccttgatlaaagccaggaacgggaagcattgc 2703  
Db 4841 ACAGAAAGTGAAGAAAGTGTGTCCAGAGAGCCCTGAGCAAAAGCCAGAGAGCCGACCTGC 4900  
Oy 2704 ctagtgtcactcagaagcctctcgaattcgaagcgaagcagattgattgatttgcac 2763  
Db 4901 ATTGTGATGCTCACCGGCTGTCCACATCCAGATGCAAGTCTTAATGTGTGTTTCAG 4960  
Oy 2764 aatggaagataaagaagaacacatcaagaagcctcctgagaatcgacaatat 2823  
Db 4961 AATGACAGAGTCAAGAGAGATGAGCAGCATGCAAGCTCTCTGCAAGAAAGCAATCAT 5020  
Oy 2824 tttaagttagttgaatgcagtcag 2848  
Db 5021 TTTTCATGATGCTAGTGTCCAGGCTG 5045

RESULT 15  
AX105057  
LOCUS AX105057 4279 bp DNA PAT 30-Apr-2001  
DEFINITION Sequence 1 from Patent WO0123540.  
ACCESSION AX105057  
VERSION AX105057.1 GI:13921209  
KEYWORDS  
SOURCE  
ORGANISM  
Canis familiaris  
dog.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 4279)  
Stocker, P.J., Steimel-Crespi, D.T., Crespi, C.L., Reif, T.C. and  
Patton, C.J.  
P-glycoproteins and uses thereof  
Patent: WO 0123540-A 1 05-Apr-2001;  
JOURNAL  
GENTEST CORPORATION (US)  
FEATURES  
source  
1..4279  
location/qualifiers  
/organism="Canis familiaris"  
/db\_xref="taxon:9615"  
17..3862  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC37718.1"  
/translation="MDPEGRKRSKAEKFNFMGRKSKNEKREKRPVSTFAMFRYSN  
WLDRLMYLVTMAALIHGAALPLMMLVGNMTDFSPANAGISRNKTPVILINESITNKT  
OHPLNHLEEMTTRAYYSIGTAVLVAAYIOVSPWCLAAAROLKIRKOFHAIMO  
EIGWEDVDHNGEILNRLPDYKINTEGIGMPEFISIAFPFGITVIRVARGKRL  
VILAIISPYLGLISAIIMAKTILSSFTDKELAYAKGAAVEVLAIRVIAFGOKKEL  
ERYNNLEAANGIGIKRAITANISIGAFLIIVASYALAEWYGSVLVSSYSIGVYL  
TVFSPVLISAFSIOASPSIEAFANARAAVEIRKIIDNPKSIDSYSGKHGPDNIG  
NLEFKNHFSYSRKEVYILKGLNKKVSGOTVALVSGSGKSTVQIMORLYDPND  
GMVICDDGDIITIVRHLREITGVVSGEPVLFATTIENINRYENVTMBIEAAYVE



ANAYDFIMKLEPNKEDTLVGERGAOLSGGOKORIAIARALVNRPKILLIDEATSLDTE  
SEAVVAVLADKARKERTTIVIAHRLSTVNRNIPFEDGVVEKGNDELMKEGIIY  
FKIVMORNGNEILEENATGSKESDLEMSPKDSGLKRESTRSIIHAPQODR  
KLGTGEDENENVPVPSFWRIKLKINSTEMPRVFGCLINGLODAPRSIIFSRIGI  
FTRDDPETKRONSMESVPLVLCIIFITPFIDOGFPFGKAGELITRIYMRSM  
LRQDSWDDPKNTGALTTRILANDAAOVKGITGSLVITVONINLGTIIISITIC  
WOLITLLAIVPIAIAIVAGVENKMLSGOALKDKLEAGKATATEINERFVSLTR  
EOKFEYMAQSIQOVPRNSLRKAHIEGVFSITQAMSFYSGCFRFGAYLVANFEM  
FODVLVSAIVFAGMAVGOSSFPADYAKASAHYMIIEKSPILDSYSPHLK  
NLTGENMPEVENVPRPOIPIVLOGISLEVKKGOTLALVSGSGSKSTVOLLERF  
YDPLAGSVLIDSKELKHLNVMRLAHLCIGVSEPIPLFCSIAENLAXGDNVSHEE  
IMQAKENAIHIFITLPEKTRIVRVDGTQISGOKORIIARALVRPPIILLIDEA  
TSADTSEKVVQALDRAREGRTCIIVIAHRLSTIONADLIVFONGVKVEHGHQOL  
LAOKGIYSMSVQAGAKR"

BASE COUNT 1294 a 834 c 1008 g 1143 t  
ORIGIN

Query Match 35.9%; Score 1026.6; DB 6; Length 4279;  
Best Local Similarity 62.3%; Pred. No. 1.6e-197;  
Matches 1699; Conservative 0; Mismatches 984; Indels 42; Gaps 4;

166 attatgataagaaccgataagataactttccacagctggatataaactgatacc 225  
1124 ATATTGACATTAACCAACCATTTGACAGCTATTGAGAGTGAACATAACAGATTAAT 1183  
226 atagaagaactgtggaattaaatgttcttccatcatcatcaagaccatcatc 285  
1184 ATTAAGGGAATTTGGAATTCAAAATGTTCACTTCAGTTACCTTTCGAAAAGAGTT 1243  
286 aagattctgaagaactgaaatcctaagaattgaatctggagagaagctgccttggtcgt 345  
1244 AAGATCTTAAGGAGTCTCAACCTGAAGGTTCAAGAGTGGGACAGAGGCGCTGTTGGG 1303  
346 ctcaatggcctgtggagaagataagtagtgcagcttctggaagaggtatgttcggat 405  
1304 AACAGTGGCTGGCGAAGACACGCTGAGTGAATGAGAGGCTCTATGAGCCACACA 1363  
406 gatggcttcatcagtgtgataagatgacatcagaagctttaaagtgcgcattatcga 465  
1364 GATGGCATGTGCTGTATTTGATGAGCAGACATTAGACATTAATGTAAGGCATCTTGG 1423  
466 gaccataatggagtgtagtcaagaagcctgttctgtcgaggaccacatcaagtaaat 525  
1424 GAATTTACTGGTGTGAGATGACAGAGCCTGTGTTGTCACACAGATAGCTGAAAC 1483  
526 atcaagaatggagagatgtgtgactgataagaagatgagaagcagaagaagaagca 585  
1484 ATTGCGTATGGCCCGGAATATGACCATGATGATGAGAAAGCTGTTAAGGAAGCC 1543  
586 aatgcgatgattatcatcagagagtttccataataatttaacatigtgtaggggaaaaa 645  
1544 AATCCCATATGATTTTATCATGAACCTACCTTAATTAATTTGACCTGTTGGAGAGGA 1603  
646 ggaagctcaaatgagtgagggcagaagaagagatcgcaattgctgtgcttgaattcga 705  
1604 GGGGCCAGCTGAGTGTGGGACAGAAACAGAGAAATCGCCATTGCTGGGCCCTGTTGCG 1663  
706 aaccccaagaattgatttgaatgagagctacgtcgtcccgagttcagaagaagaatca 765  
1664 AACCCCAAGATCTTTCGCTGGATGAGGACAGTCAAGCTTCGACACTGAAAGTGAAGCA 1723  
766 gcttctcaagctgacagtgagaagcagcaagaagtcgagatacaatcgtgtgtagcacac 825  
1724 GTGGTTACAGTGGCCCTGAGTAAGGCCAGAAAGCCGAGATACCATGTGTATAGCTAT 1783  
826 cgaattctactatcgaagtgcagattgattgtgacccataaagatggaatgcgtgcg 885  
1784 CGTTGTCTACAGTTGCGTAATGCCGATCATGTCTGTTTGTGATGATGAGGATATTGTG 1843  
886 gagaagaagacacatgctgaagataatgcgaagaagagtcataatctactgtgtg 945  
1844 GAGAAAGGAATCATGATGAACCTCATGAAGAAGAGGCAATTTACTTCAAACTGTGACA 1903

946 tcacaggatataaaagaactgataac- agatgagtgcaatgacatattctactgaaag 1004  
1904 ATGCAGACACAAGAGAAATGAATTTGAGTGAATAATGCACTGGTGAATCCAAAATGTGA 1963  
1005 aaagaccacactcactccctgcgacactgtgaaag- -----catcaagtcagacttc 1056  
1964 AGTAGAGCTTGGAAATGTTCTCCAAAAGATTACAGGCTCCAGTTTAATTAAGAAAGATCA 2023  
1057 attgacaagcgtgaggaatccaccacatcctaagaagataagtc- ----- 1101  
2024 ACTCGAGAGGATTTACATGACCAAGCCAGACAGACAAAGCTTGCTCAAAAGAGAC 2083  
1102 -----cctgaagctctcattataaaatttaaaagttaacaagcctgaa 1146  
2084 TTGAATGAGAAATGTACCTCCAGTTTCTCTGAGGATTTCTGAAGCTGAACTCACTGAA 2143  
1147 tggccttggtgtctctggagacatcggctcctgctcctaagtgaactgtcatccagta 1206  
2144 TGCCCTTATTTTGTGTTGTTGATTTTGTGCTATTATTAACGAGAGCCCTGCAACAGCA 2203  
1207 ttctccatcactcttgcaaaaattataacacatgttg- --gaaataatgataaaccaca 1263  
2204 TTTTCAATATATATTTCAAGGATTTATAGGATCTTTAACCGAGATGAGATCCTGAACA 2263  
1264 ttaagaatgatacagaatattatccatgatatcgtcaatttggtgtatctgtt 1323  
2264 AAACGACAGAAATGTAACATGTTTGTGATTTCTTACTGCTTGAATTTATTTCTTT 2323  
1324 gtacgtattctatcagaggaattatttaacgagagaaggggaaatttaacatgaga 1383  
2324 ATTTCATATTTTCTCCAGGCTTCACATTTGGCAAACTGGGAGATCCCTCAATAGCGG 2383  
1384 ttaagaacttggccttcaagaagcattatcagaatattgacctgtgtgttgatgaag 1443  
2384 CTTCGATACATGTTTCAATCATCTGATGACAGATGCTGATGCTGTTGATGACCT 2443  
1444 gaaacagcagaagagcttgcagaacaaatttgccatagatataagcaaaattcaaga 1503  
2444 AAAAACACACTGGAGCATTTGACAAACAGGCTTCCATATGATGCGGTCAAGTTAAAGG 2503  
1504 gcaacaggttccagagatttgccttcaacaaatgcaactcaacatcagatcctcagtt 1563  
2504 GCTATAGGTTCCAGGCTTGCCTGCTATTAACCAAGATTAACCAAACTTGGGACAGCAT 2563  
1564 atcattccttataatagatgagagagatgacatccatccatctgactgagatgtctcagta 1623  
2564 ATTATATCTTAATCTATGTTGCAATTAACACTTTTACTCTTACGCAATTTGACCATC 2623  
1624 ctgctcgtagaagaatgattgaaacccgacagcaatgactgattgccaacaagaataag 1683  
2624 ATTGCAATAGCAGGAGTTGTTGAATGAAATGTTTGTCTGGACAAAGCATGAAAGATAG 2683  
1684 caagaacttaagcgtctggaagaatagcaactgaaagcttggagaatatacgtactata 1743  
2684 AAAAGCTAGAGAGAGCTGGGAGATTTGCTACAGAGCATGAAACTTCCGAACGTGT 2743  
1744 gttcatthaacaaggaanaagccttcgagcaaatgatatgaagagatgcttgagctcaa 1803  
2744 GTTCTTTTGAAGTCCGAGCAAAAGTTGAATGATGATGATGACAGAGATTTGCAAGTACA 2803  
1804 cacaagaataactcgaagaagaacagatatatttgaagcgttctgcatcagaacagtc 1863  
2804 TACAGAAACTTTTGAGGAAAGCACACTTTCGGGCTCATTTTCTATATCCACAGCA 2863  
1864 ttatataatttgcctatgacagcaggttctgatttggagcctatthaatcaagctgga 1923  
2864 ATGATGATATTTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2923  
1924 cgaatgaccccaagagagatgtcatagtttcaatcagatcagatcagatgagc 1983  
2924 TTCAATGAATTTCAAGAGATGTTCTTTGTTGATTTCAAGTATTTGCTTGTGTCATATGCA 2983





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:37:44 ; Search time 293.81 Seconds  
(without alignments)  
2201.495 Million cell updates/sec

Title: US-09-873-409-10

Perfect score: 2856  
Sequence: 1 cctaattcctctaatatctc.....atgcacagtcagtcagtgta 2856

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1030.2	36.1	4646	1 US-08-181-471-2	Sequence 2, Appli
2	1030.2	36.1	4669	6 5206352-3	Patent No. 5206352
3	1030.2	36.1	6505	2 US-08-793-610-5	Sequence 5, Appli
4	1030.2	36.1	9318	2 US-08-793-610-6	Sequence 6, Appli
5	1028.6	36.0	4264	2 US-08-784-649A-1	Sequence 1, Appli
6	1028.6	36.0	4264	2 US-08-784-649A-5	Sequence 5, Appli
7	1025.4	35.9	4669	2 US-08-583-276-18	Sequence 18, Appli
8	1022.2	35.8	4669	2 US-08-752-447-1	Sequence 1, Appli
9	953.2	33.4	4233	3 US-09-120-513-1	Sequence 1, Appli
10	953.2	33.4	4233	4 US-09-450-105-1	Sequence 1, Appli
11	856	30.0	2726	1 US-08-461-823-1	Sequence 1, Appli
12	414.6	14.5	4002	2 US-08-996-545-1	Sequence 1, Appli
13	414.6	14.5	4002	2 US-08-996-545-3	Sequence 3, Appli
14	414.6	14.5	4002	2 US-09-328-320-1	Sequence 3, Appli
15	414.6	14.5	4002	4 US-09-328-320-3	Sequence 3, Appli
16	374	13.1	4047	2 US-08-612-734B-1	Sequence 1, Appli
17	309.2	10.8	4224	1 US-08-612-521-1	Sequence 1, Appli
18	302	10.6	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	283.2	9.9	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	196	6.9	3927	2 US-08-996-644-1	Sequence 1, Appli
21	196	6.9	3927	3 US-09-352-552-1	Sequence 1, Appli
22	195.2	6.8	3924	2 US-08-996-644-3	Sequence 3, Appli
23	195.2	6.8	3924	3 US-09-352-552-3	Sequence 3, Appli
24	181.6	6.4	3909	1 US-08-232-537-1	Sequence 1, Appli
25	149	5.2	2061	4 US-09-061-764A-17	Sequence 17, Appli
26	145.2	5.1	2407	2 US-08-895-522-2	Sequence 2, Appli
27	145.2	5.1	2407	3 US-09-195-391-2	Sequence 2, Appli

28	145	5.1	5120	3 US-08-772-270A-6	Sequence 6, Appli
29	145	5.1	8370	2 US-08-488-706-1	Sequence 1, Appli
30	144.8	5.1	7721	3 US-08-772-270A-14	Sequence 14, Appli
31	141.2	4.9	1959	4 US-09-061-764A-4	Sequence 4, Appli
32	133.4	4.7	1201	3 US-08-961-083-189	Sequence 189, App
33	133	4.7	2244	4 US-09-061-764A-18	Sequence 2, Appli
34	130.4	4.6	4403765	4 US-09-103-840A-2	Sequence 2, Appli
35	128.6	4.5	2376	1 US-08-394-880B-1	Sequence 1, Appli
36	128	4.5	4403765	4 US-09-103-840A-2	Sequence 2, Appli
37	128	4.5	4411529	4 US-09-103-840A-1	Sequence 1, Appli
38	117.6	4.1	6143	1 US-08-612-521-3	Sequence 3, Appli
39	110	3.9	28804	2 US-08-592-874-1	Sequence 2, Appli
40	110	3.9	28804	3 US-09-096-942-2	Sequence 2, Appli
41	110	3.9	28804	3 US-09-096-867-2	Sequence 2, Appli
42	104.2	3.6	5889	1 US-08-463-092B-5	Sequence 5, Appli
43	104.2	3.6	5889	2 US-08-462-109A-5	Sequence 5, Appli
44	104.2	3.6	5889	2 US-08-460-907B-5	Sequence 5, Appli
45	104.2	3.6	5889	3 US-08-463-179A-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-08-181-471-2  
Sequence 2, Application US/08181471  
Patent No. 5641508  
GENERAL INFORMATION:  
APPLICANT: Li, Lingna  
APPLICANT: Lishko, Valery K.  
TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Thomas Fitting  
STREET: 12526 High Bluff Drive, Suite 300  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92130  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,471  
FILING DATE: 13-JAN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: ANT0029P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-792-3680  
TELEFAX: 619-792-8477  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4267  
US-08-181-471-2



Db	3626	CTGCTGGGCAAGCACTGGCTTGTGGGAAGACACAGTGTCCACTCTCTGGAGCGTTTAC	3685
QY	2284	gacccccgtgaagaagaacagtctgttttga tgg tgtgtaatgcaaaagaa ttgaatgtaag	2343
Db	3686	GACCCCTTGGCAAGGAAAGCTGCTTTGATGCGAAAGAAATTAACGACTGAATGTTACG	3745
QY	2344	tgagtcggttccccaaataagaatcgttctctcaagagccctgctcttcaacttgaagact	2403
Db	3746	TGGCTCGAGACACCTGGGCAATCGTGTCCAGAGAGCCCATCTCTTTGACTGCGACGCTT	3805
QY	2404	gcttgaagacac tgcctatgtgtgaacaacagcgt tgtgtgtcca ttgaatgtaatcaaaagaa	2463
Db	3806	GCTGAGAACTTGGCTATGTGAGACAAACAGCGGGGTGTGTACAGAGATGCTGTAGG	3865
QY	2464	gcgcgaatgcaggaatatccattcttatttgaaggtcttccctggaataatacaacaa	2523
Db	3866	GCACCAAAAGAGGCCAATCAATATGCTTATCATGATCATCTCTCTTAATTAATATATGACCT	3925
QY	2524	caagtctgagcttgaagaagagacacagctcttctgcgcgcgaagaacaagaactagtatttga	2583
Db	3926	AAATAGAGACAAAGGAAGACTGACTGTCTGTGGGGCCAAACAAACGATTTGCCATAGCT	3985
QY	2584	aggagctcttctccaaaaccccaaatattatgttggatgagggccaacttcagcccttgat	2643
Db	3986	CGTCCCTCTTGTATGACACGCTCTCATATTTTGTCTTTTGGATGAAGACCACTGACGCTTGAT	4045
QY	2644	aatjacagttagaaggtgtgtctgacatgccttgataaagacgaagacggaagaagacagc	2703
Db	4046	ACGAAAGTGAAMAAGTTGTCCAAAGAAAGCCCTGTGAGCAAAAGACGAGAAAGGCCGACCTGC	4105
QY	2704	ctagtggttcaactcaacaggtctcttgcgaatttcgaacgcagatttgaatgtgtgtctcac	2763
Db	4106	ATTTGTGATGCTCAACCCCTGTCCACCATCCAAAGCAAGCACTTAATATGTGTGTTTACG	4165
QY	2764	aatgagaagataaagaagaacgaagacttcaagagctctcttgaagaatcogagacatatat	2823
Db	4166	AATGGCAGAGTCAAGAGAGCATGTGGCAGCATCAACACACTCTGTGGCACAGAAAGCATTCAT	4225
QY	2824	tttaagttagtgtaatgcacagctaacg 2888	
Db	4226	TTTTCAATGTCAAGTGTCCAGCGCTG 4250	

```

RESULT 2
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLOVES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
; SEQ ID NO:3
; LENGTH: 4669
5206352-3

```

	Query Match	Similarity	36.1%	Score	1030.2	DB	6	Length	4669;
	Best Local	Similarity	62.2%	Pred.	Mism.	1.4e-286;			
	Matches	1995:	Conservative	0;	No.Matches	986;		Indels	42; Gaps
Oy	166	attactgataagaaccccgatataagataaccttcttcaccacgcctggatataaactgtatccc	225						
Dd	1526	aataattgatataaatgaoccaagtattttacaacattcttcgaagaatggcagcacaaaaggtaatt	1585						

QY	226	atagaaggaactgtggaatttaaagtgttcttctaattacatccaagaccatctac	285
Db	1586	atlaagaggaataatttggaatttcagaaatggttccattccagtaccatctcgaaaaagaatc	1645
QY	286	aagattccgaagaaggtctgaattccagaattaaagtcctgagagacaagctgcgttcgtgc	345
Db	1646	aagatcttgaaagggcccggaaccctggaagagtcgcaagtcggacagccggtgcctgtgtga	1705
QY	346	ctcaatgycgaatgagaaagatlacgtgtacagcttcctcttcgacaggttatatccggat	405
Db	1706	aacagtcgctgtgagaaagacacaacagctccagctgatgcagaggtctatgaaccaca	1765
QY	406	gatygccttalcatagtgtgatgagaatgacatcaagagctttaaagtctgcgcattatcga	465
Db	1766	gagggagatgctcaagtctgttatgacagaagatattatagaccataatgtlaaggtttctacg	1825
QY	466	gacctattggaagtggtttagttcaagagccgtttgtctcggaagccccaatcgaatacat	525
Db	1826	gaatcatctgtgtgtgtgagtcgaagaccctgttatgtctgcaccctgaatctgaaaac	1885
QY	526	atcaagtatgacgagatlgatgttactcgtatgaagaaatgagagagcaagaagaaaga	585
Db	1886	attcgctatgcccgttggaataatgacacatgatgatgagattggaagacgttcaaggaagcc	1945
QY	586	aatgcgtatgatttcatcatgagattccctaataattaatcatatcatgtgttagggaaaaa	645
Db	1946	aatgcctatgacttcatcatcagaactcgtccataaatttgaccacccctgtttgagagaga	2005
QY	646	ggagctcaaatgaaatlgagagggcagaaacaaagagatctgcgaattgtctgttccttagtga	705
Db	2006	ggggcccgatgtgagttgtgtgagaaagcaagagatctgcgaattgtcaggtccctgtctgc	2065
QY	706	aaccccaagattccgatttttagatgaagagcagctgtccctgtgattcagaagaagcaatga	765
Db	2066	aaccccaagatccctcccgctgagatgagggccagctgcagctcttgaccagaaagcgaaagca	2125
QY	766	gctgttcaagctgtgacctgagaaagcgagacaaagctcggactacaalctgtgtgacaac	825
Db	2126	gtgttctcaagtgctctcttgatgaagccagaaagctcggaccacacattgtgtatgattcat	2185
QY	826	cgacttctactatctgaagtgcgaattttgatgtgacacctaaagatgaaatgctgtgcg	885
Db	2186	cgttgtctacagttctglaatgctgcagacgtatcatgcgttgcgtatgatatgagatcatgtg	2245
QY	886	gagaaagagacacatgctgcgaactaattgcgaaacggagctctatattcattacatctgtgtg	945
Db	2246	gagaaaggaataatcatgatgaacatcagaagaaagaaagagcattacttcaactgttcaca	2305
QY	946	tcaacagatatataaanaagctgtgaatgaacagatgagatcacaatgacatattca-----	997
Db	2306	atgcagacagcaggaagaaatgaagtgtgaattagaanaatgcagctgatgataccaaagtga	2365
QY	998	-----ctgaaagaaagaccacatctctctctgcactctgtgaaagagacalca---	1045
Db	2366	attgatctctgtgaaatgcttctcaaatgattcaagatccagttcaatagaanaaagatca	2425
QY	1046	-----agtcagacttattgacaagcgtgaggaatccaccatct	1086
Db	2426	actcgttagtgatgtccgtgtgatacacaagcccaagacagaaagcttgcaccaagagct	2485
QY	1087	aaagagataagctctctcgtgaagtcctctcataltaanaattttaaagttaacaacagctgaa	1146
Db	2486	ctgagtaagaatatacctccacagtttctcttcttgagagattatgagagtaaatattacatgaa	2545
QY	1147	tgagctttgtgtctcggggagacttgctctcgttcttaagtgaactgttataccagta	1205
Db	2546	tgagccttattgtgtgtgtgtatatttgycacattaaatgagagccctgacacacgca	2605
QY	1207	tttccacatcttcttgcaaaatatataacatgtttggaaataatgataaaac---aca	1265
Db	2606	tttgcaataaataatttccaagaatatataggggtttttcaacaagaattgagatccctgaaaca	2665
QY	1264	ttaaagcatgtatgcagaataatttcatcatgatalctgcgtcaatttgggttattgtctt	1325

5237

Db 2666 aacagacagatagtaactgttttcaactatgtttccagcccttggaattattcttt 2725  
Oy 1324 gtcagttatttcaagcaggatattttaagccagacaggggaaattttaagataga 1333  
Db 2726 attcaattttccctccaggtttcacatttgcgaaagcttgagagatccctcaacacg 2785  
Oy 1384 ttaagacattgtgcttcaagacatgttatatcagatattgcctgttgatgaaag 1443  
Db 2786 ctccagatagatgtttccagatccatgtctcagacagatgtgagattgtttgagccct 2845  
Oy 1444 gaaacacagcagagagctgtgacacaaatattagccatagatagcaaaatcaagga 1503  
Db 2846 aaaaacacacagcagatgtactaccagagctgcgcaatgtgtcgtcaagttaaag 2905  
Oy 1504 gcaacagttccagagattgtgcttcaacaaatagcaactaacataggagcttcgtt 1563  
Db 2906 gctataggttccaggtctgtcgtaattacccaagatagaaatcttgagacagata 2965  
Oy 1564 atcattcccttataatagatgagagatgacatccctgtatctgtatgtccagta 1623  
Db 2966 attatcccttcatctatgtgtgcaactaacactgttacccttagcaattgtaccatc 3025  
Oy 1624 ctgctcgtgacaggaatgtatgaaacccgacgaatgtacgtattgccaacaaagaa 1683  
Db 3026 attgcaatagcagagatgtgtgaaatgtgtctgtgacaaagcactgaaagataag 3085  
Oy 1684 caaaatattgaagctgtgagaaagatagcaactgaagcttgagaaatatacgtata 1743  
Db 3086 aaaaagacagagagctgcgggaaagatcgctactgaaagaaatagaaatcccgaa 3145  
Oy 1744 gtgcattaaacagaggaagacccctcgagcaaatgtatgaagagatgcttcaactaa 1803  
Db 3146 gtctctttagactcagagcagaaagtltgaacatagtatgtctcagagatttcagaa 3205  
Oy 1804 cacaagaataccctcgaaagacacagatatgtgaaagctgttatgcaatcagatgcc 1863  
Db 3206 tacagaacactcttgagaaagacacacactcttgaaatattccttccacacagca 3265  
Oy 1864 ttatatatttgcctatgcagcagggtttgattttgagccttattcaagcttga 1923  
Db 3266 atgatgtatttttccctagctgcgagtgcttccggtlttgagcctactgtgtgcataa 3325  
Oy 1924 cgaatgaccccaagggagcatgtcatagtttctacgaattgataagagacatagcc 1983  
Db 3326 ctcatgagcttgagagatgtctgttagtatttccagctgtgtcttgagcagatgcc 3385  
Oy 1984 atcggaacaaacgctcgtttgtgctcctgaatattccaaagccaaatcggggctgc 2043  
Db 3386 gtggggcagatcagttcatltgtcctcactatgcgcaagccaaatatacagccac 3445  
Oy 2044 ctgttgcctgttggaagaaacaaatataagacacgacgcatcaagagggaaaaag 2103  
Db 3446 atcatalgacatltgaaaaaaccccttgatgtacagacatcaacagagagcctaag 3505  
Oy 2104 ccagacacatgtgaaggaatttagagatttcgaagaagctcttctcttaccatgtgc 2163  
Db 3506 ccgaacacatgtgaaggaatgtcaacattgtgaagltgataactatccacccga 3565  
Oy 2164 ccagatgttttcatcctcgtgtgcttaccctcagatattgacgagaaagacgttag 2223  
Db 3566 ccggaatcccgatgtctcagagactgtgagctgtgagtgaaagagggcagagactgc 3625  
Oy 2224 ttgttgaggagcagcgtgtgaggaaagcactctgttcaacttctcagagacttat 2283  
Db 3626 ctgggtggcagcagtgctgtgagaaagcagtgctcagctctcctcagacggtttcac 3685  
Oy 2284 gaccccggtgacaggaagatctgtttgattgtgtgagtgcaaaagaaatlaagttacag 2343  
Db 3686 gaccccttgagaggaagtgctgtctgagtgagcaaaagaaatgaagcagctgaagttcag 3745  
Oy 2344 tgcctcgttcccaatagcaatcgttctcaagagcctgtgtcttcaactgacgact 2403

Db 3746 tgcctcagacacacctgggcatcgtgtccagagagcccatcctgtttgactgcagact 3805  
Oy 2404 gctgagaacatcgccctatgtgtgacaacagccgtgtgtgctcattagatgaaatcaagaa 2463  
Db 3806 gctgagaacatcgccctatgtgtgacaacagccgtgtgtgtgtcacaagagatcgtgag 3865  
Oy 2464 gcccaatgacgaatatalcatcttatttgaaggtctccctgagaaatatacacaa 2523  
Db 3866 gcagcaagagggcacaacatcatgtccttcatcgatcgtcgttaataatataagact 3925  
Oy 2524 caagtgtgactgaaagagcagacttctgtgcgccaagaaacaaagactcattgtca 2583  
Db 3926 aaagttagagacaaggaactcagctctgtgtgccaagaaacaaagcattgcatagct 3985  
Oy 2584 aggcctcttcccaaaaacccaattttatgtgtgagtgagggcaccatccgctcgat 2643  
Db 3986 cgtgccttgttgaacagcccatatttggttttgattgtagagcaagtcagctctgat 4045  
Oy 2644 aatgacagtgagaaagtgtgtcagcatgtcccttgataaagccagagcaggaagacatgc 2703  
Db 4046 acagaagtgaaaaagtgcttccagaagcccttgacaaagcagagagagcgcgacctgc 4105  
Oy 2704 ctatgttactcaacagctctctgcgaattcagaagcagattgtatagtggtctgcac 2763  
Db 4106 attgtatgtctcagccgtgttccacatccagaatgcagacttaactaattgtgttccag 4165  
Oy 2764 aatgaaagataaaggaacgaagactcatcaagagctcctgagaaatcgagacataat 2823  
Db 4166 aatgacagatgcaagggacatgagcagcagcatcagcagctgtgcgcaacgaaagcactat 4225  
Oy 2824 ttaagttagtgaatgacagtlcag 2848  
Db 4226 ttccaatgtcaggtgtccaggtcg 4250

RESULT 3  
US-08-793-610-5  
Sequence 5, Application US/08793610  
Patent No. 5858744  
GENERAL INFORMATION:  
APPLICANT: BAUM, Christopher  
APPLICANT: STOCKING-HARBERS, Carol  
APPLICANT: OSTERIAG, Wolfram  
TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
TITLE OF INVENTION: FOR GENE TRANSFER  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/793,610  
FILING DATE: 07-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 31 973.8  
FILING DATE: 08-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 195 03 952.1  
FILING DATE: 07-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP95/03175  
FILING DATE: 10-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Berman, Richard J.  
REGISTRATION NUMBER: 39,105



OY	2044	ctgttgccctctgttggaagaagaacccaatatagacaagccgagctcaaggaagggaag	2103
Db	4838	ATCATCATGATCATTTGAAAAAACCCTTTGATTGACAGCTACAGCAGGAAAGCCTAATG	4837
OY	2104	ccagacacatgttgaagggaatttagagtttcgagaagtcctcttcctatccatgctgc	2163
OY	2164	ccagatgttttccatcctcgttggtcttatccctcagtatattgagcgaagaagacagtagca	2223
Db	4958	CCGACATCCCACTGCTTCAGGAGCATGAGCCTTGAGAGTGAAGAGGCGACAGCGCTGAGCT	5017
OY	2224	tttggggagagcagcggcgtctgggaaaaagcacttcctgttcaactcttcgagagcttat	2283
Db	5018	CTGTGTGGGCAAGCACTGGCTCTTGGAAGACACAGTGTCTCACTCTCTTGAGGGTTCTAC	5077
OY	2284	gaccgccgtgcaagagcaagctgctctttagtgytgytgcatacaaaagaaattgaatgcacg	2343
Db	5078	GACCCCTTGCGAGGAGAAAGTGCCTTGATGGCAAGAAATTAACGCACTGAATTTTCAG	5137
OY	2344	tgagctcggttcccaaatagcaatcgttccctcaagagcctgtgctcttcaactgcagcat	2403
Db	5138	TGGCTCCGAGACACCTGGGCACTCGTGTCCAGAGACCCTCATCTGTTTGACTGCAGCATY	5197
OY	2404	gctgaagaacatcgctctatggtgtgcaaacagcgcgtgtggtgtccattagatgagatacaagaa	2463
Db	5198	GCTATGAGACATTTGCTTATGGAGAACACACCGGGTGTCTCACAGGAAGAGATCCTGAGC	5257
OY	2464	gcccgaatgtgaagcaaatatccactcttattatgaagctcctccctgagaatatacaaca	2523
Db	5258	GCACCAAGAGAGGCGCAACATACATGGCTTCATCGACTCCTCCTAATTAATATGACACT	5317
OY	2524	caagttggaactgaagaaggacagcgttctcbygcgcgaagaacaaagactagctatgca	2583
Db	5318	AAATAGGAGAGCAAGAGCAATCACTGCTCTGTGGTGCCAAACAGCATTGCCATAGCT	5377
OY	2584	aggcgtcttcccaaaaacccaataatttatattgttgaatgaagccacttcagcctcgat	2643
Db	5378	CGTCCCTTGTATGACAGCCTCATATTATTTTGCTTTTGATGGAAGCCACAGCTCAGCTCTGGAT	5437
OY	2644	aatcacagttgagaagcgtggtctgacgaatcccttgaataagcagcagcagggaagacatgc	2703
Db	5438	ACAGAAAGTGAAAAGGTTGTTCAGAAAGCCCTTGAGACAAAGCCAGAGAGGCCGACCTTGC	5497
OY	2704	ctagtggtcaactcaagcgcctctgcaatlcagaacgcagattgataagttggttctgac	2763
Db	5498	ATTGTGATTGCTACCCGCTGTCCACATCCAGAAATGACAGACTTAATAGTGTGTTTAG	5557
OY	2764	aatggaagaataaaggaagaagaactcatcaagagctcctctgagaatctgacatatat	2823
Db	5558	AATGGCAGAGTCAAGAGGACATGTGCAGCAGCAGCACTCTGTGGCAGAGAAAGCATCTAT	5617
OY	2824	tttaagttagtgaatgcagcagtagc	2848
Db	5618	TTTTCAATGTCAGTGTCCAGGCTG	5642
RESULT 4			
US-08-793-610-6			
Sequence 6, Application US/08793610			
Patent No. 5858744			
GENERAL INFORMATION:			
APPLICANT: BAUM, Christopher			
APPLICANT: STOCKING-HARBERS, Carol			
APPLICANT: OSTERFAG, Wolfram			
TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF			
NUMBER OF SEQUENCES: 6			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Nikaido, Marmelsteind, Murray & Oram LLP			
STREET: 655 Fifteenth Street N.W. Suite 330			
CITY: Washington			
STATE: D.C.			

COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/793,610  
FILING DATE: 07-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 31 973.8  
FILING DATE: 08-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 195 03 952.1  
FILING DATE: 07-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP95/03175  
FILING DATE: 10-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Berman, Richard J.  
REGISTRATION NUMBER: 39,105  
REFERENCE/DOCKET NUMBER: P1614-7007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)638-5000  
TELEFAX: (202)638-4810  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9318 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA  
US-08-793-610-6

[illegible]

[illegible]

D	4437	AAAGACTAGCAAGTCTGCTGGGAAGATCGCTACTGAAAGCAATGAAAACTTCCGAACGGTT	4496
Q	1744	gtgcatacaacaaggagaaaaagcccttcagagcaaatgataagagagaagcttcagactcaa	1803
D	4497	GTTCCTTTGACTCGAGAGCGAAGATTGGAACATATGTATGTGTGCAGAGTTTCACAGTACCA	4556
Q	1804	cacagaataatcctcgaaagaagacagatatatgtgaagctgtgatacattcaagccatgcc	1863
D	4557	TACAGAAACTCTTTGAGAGAAAGCACATCTTTGGAAATTAACATTTTCCCTTACCACGACA	4616
Q	1864	tttatataatttgccctatcagcagaaggttcgaatttgagagccatttaattcaagctcga	1923
D	4617	ATGATGTAATTTTCTATAGCTGATGATGTTTCCGTTTGAGCCCTACTTGCTGGCACAATAA	4676
Q	1924	cgaaatgaccccaagaggacatgcttcatagitttttaäctgcataatgcataatgagatagcc	1983
D	4677	CTCATGAGCTTTAGAGATGTGTTGTTAGTATTTTTCAGCTGTTGTCTTGGTCCATGGCC	4736
Q	1984	atcggaanaacgctcglttttgctcctgaatatccaaagccaatcggggctgycgat	2043
D	4737	GTGGGGCAAGTCATTTATTTGCTGCTGACTGATGCCAAAGCCAAAATATACGACACCAC	4796
Q	2044	ctgttgccctgttggaagaagaacaataatagagaagccgagcaagaaggagaaaag	2103
D	4797	ATCATCATGATCATTTGAAAAAACCCCTTGATTTGACAGGTACAGACGAGAGGCTTAATG	4856
Q	2104	ccagacacatgtgaaggaaatttagatttcgagaagcttccttcctatccatcatalgtgc	2163
D	4857	CCGAACACATTTGGAAGAAATGTACATTTGGTGAAGTTGTATTCACATCCACCAGA	4916
Q	2164	ccaagatgtttcaatccctccgttggtctatccctcagtatatgagcgaagaagacagtaga	2223
D	4917	CCGGACATCCCACTGCTTTCAGAGGACACTGAGCCTGAGGTAAAGGGCCAGACGCTGGCT	4976
Q	2224	tttgggggagacagcgctcttggaagaagaccttcgttcaacttcctcgagaagacttat	2283
D	4977	CTGGGGGACACACTGGCTGTGGGAAGACACAGTGTCTCAGCTCTGTGAGCGGTTCTAC	5036
Q	2284	gaaccccgycagaaggacaagctgcttctgaatggtgtgagatgcacaaagaatttgatgtaag	2343
D	5037	GACCCCTTGCGCAGGGAAGTGTGCTTGATGCGCAAGAAATTAAGGCACTGAAATGTTGAC	5096
Q	2344	tggtcgcgttcccaataagaatcgttcctcaagagcctgctcctcaactgcagcaat	2403
D	5097	TGGCTCCGACACACCTGGCATTGTGTCCAGAGACCCATCCTGTTTGACTCGACGACTT	5156
Q	2404	gctgagaacacacgccttatggtgacaaagccgtgtgtgtccattgatatgagataaaga	2463
D	5157	GCTGGAACATTTGCTTATGGAGAAACAGCGGGGTGTCTCACAGGAAGATGCTGTAGG	5216
Q	2464	gcccgaatgtgcgaataatccattccttatggaagctctccctcgtagaataatacaaca	2523
D	5217	GCAGCAAAAGGAGGCCAATACATACCTTATGATGATGATCAGTGCCTATATTAATATGACT	5276
Q	2524	caagttgcactgaaggagacagcttctgcgcgcagaacaagaagactagctatgca	2583
D	5277	AAAGTAGAGACAAAGGAACATCACTCTGTGGGCCCAAAACAACGCAATTGCCATACCT	5336
Q	2584	agggtccttcccaaaaaaccaaaatttatgtgtgatatgagagccaattcagccctgat	2643
D	5337	CGTGCCCTTTGTTAGACAGCCCTCATATTTTCTTTTGATGAAGACCACGTCACCTCTGGAT	5396
Q	2644	aatgacagtgaagaagtgtgttcagcatgcacctgataaagcagagacgaggaagacagc	2703
D	5397	ACAGAAATGTAAAAGTTGTCTCAAGAAAGCCCTTGACAAACGAGAAAGGCCGACACTGC	5456
Q	2704	ctagtgctactccagagctctcgcgaatcagaacgcagatttgatagtggtctgac	2763
D	5457	ATTGTATTTGCTCACCGCGTGTCCACATCATCAAGAAATGACAGACTTAATGTGTGTTTACG	5516
Q	2764	aatgaaagataaagaagaagaactcattcaagagctcctcgtagaatctcgagactatat	2823
D	5517	AATGCCAATGTCAAGGAGCATGTGGACCATACACACTCTGTGGCAGAAAGCATATAT	5576







286 aagatctgaagaagctcgaatctcagaatlaagctcgaagaagcagtcgccttgctcgt 345  
1360 AAGATCTTGAAGGCGCTGAACCTGAAGGTGCAGAGTGGGACAGCGGCGCTGGTTGGA 1419  
346 ctcaatgcaagcaggaagtagcagtagtccagcttcgcaaggtatataatgcgat 405  
1420 AACAGTGGCTGTGGGAAGACACACACCTGATGAGAGGCTGTGACCCCA 1479  
406 gtagccttatacgtgtgagtagaagtagacatcagagcttaaatgtcgccattatcga 465  
1480 GAGGGAGTGGCTGCTGTGATGACAGAGATTAATAGCCATAATGTAAGTTTCTACGG 1539  
466 gaccatactgaagtggttgaatcagaagcctgttctgtcggagcaaccacagtaaacat 525  
1540 GAATATATTTGGTGGGAGTCAAGAACCTGTATTTTGGCCACACAGATAGTGAAGAAC 1599  
526 atcaagtagaagcagatgtagtgcagtagaagagatggaagagcagaaggaagca 585  
1600 ATTGCGTATGGCGTGAATAATGTCACCATGATGATGATGAGAAAGCTGTCAAGGAAGCC 1659  
586 aatgcgatattatcatcagagagttccctaataattatatacatgtgtaggggaaaaa 645  
1660 AATGCTATGACTTTTATCATGAAGAACTGCTCATTAATTGACCCCTGGTTGGAGAGA 1719  
646 ggaagctcaaatgagtggaagggcagaacacagaagatcgcaattgctgctgactga 705  
1720 GGGGCCCCAGTTGAGTGGTGGGACAGAACAGAGATGCCANTTGCAGCTGCCCTGTTGCG 1779  
706 aaccacaagattctgaatttaagatgagctcagctgcgcctcgagttcaagaagaagatca 765  
1780 AACCCCAAGATCCCTGCTGATGAGAGGCCACGTCACGCTTTGACACAGAAAGGAAGCA 1839  
766 gctgtcaactgacagcagcaggaagcagcagaagctcgacatacagctgtgtagcacac 825  
1840 GTGCTTAGGTGGCTCTGGATTAAGCCAGAAAGGTGCGACACCATTTGTATAGCTCAT 1899  
826 cgaattctactatctcgaagtgcagattgtagtgcacccaagaagatgagaaatgcgcg 885  
1900 CGTTTGTCTACAGTTGCTGAATGCTGACGATCMTCCGTGGTTGATGATGAGTATTTGTG 1959  
886 gagaagaagcagacatgctgaactaagcagaagagctatatactatctcagtgatg 945  
1960 GAGAAAGAAATCATGATGAAGCTCATGAAGAAAGAAAGCAATTTACTTCAAACTTGTCACA 2019  
946 tcaagagatataaagaagcagcagaagatggaagtagaagacatattca----- 997  
2020 ATGCAGACAGCAGAAATGAAGTTGAATTAAGAAATGCACTGATGATCCAAAAGTGA 2079  
998 -----ctgaagaagaagcaactcaactcctctgcactctgtgagagcatca--- 1045  
2080 ATTGATGCTTGAAGAAATGCTTCAAAATGATTCAGATCCAGCTCAATGAAGAAAAAGATCA 2139  
1046 -----agtcagactcatgacaaagctgagaaagctggaatccacccaatct 1086  
2140 ACTCGTAGAGTAGTCCGTGATGACAAAGCCCAAGACAGAAAGCTTAGTACCAAAAGGCT 2199  
1087 aagaagataagctctccgaagctctctatataaatttaagaagtaaacagaagcctga 1146  
2200 CTGGAATGAATATATACCTCCAGTTTCTTTGGAGATTAATGAAGCTAAATTTAAGTGA 2259  
1147 tggccttttgctgtctcggagacatggtcctgtcttaaatggaagcgttcatccagta 1206  
2260 TGGCCTTTTGTGTTGTTGTTGATTTTGTGCATTAATTAATGAGGCGCTGCAACACAGA 2319  
1207 ttcttcacatccttgcagaataataacatgcttgaataataagataaaaaac---aga 1263  
2320 TTTGCAATAAATAATTTTCAAAATATATAGGGGTTTATACAGATGATGATCCTGAAACA 2379  
1264 ttaagagctgtgcagaataattatccatgatatcgtcatcttggtgttattgctt 1323  
2380 AAACGACAGAAATAGTAACTTGTTCACATATTGTTTCTAGCCCTTGGAATTTATTTCTTTT 2439

1324 gtcagttattcatgcaaggtattatttaagcgagaagcgagggaataattcaagatgaga 1383  
2440 ATTACATTTTTCCTTACAGGGTTTCACATTTTGGCAAAAGCTGGAAGATCTTACCACAGCG 2499  
1384 ttaagacacttgccctcagaagcatgtlatcagaagatattgcctgtttagtgaaga 1443  
2500 CTCGGATACATGGTTTCCATTCATTCATCCAGACAGATGATGATGATTTGATGACCT 2559  
1444 gaaacagcacaagagctgtgaacaaatattagccaatagatatagcaaatattcaaga 1503  
2560 AAAACACACACTGGAGCATGTGACTACACAGGCTGCCAATATGCTGCTCAAGTTAAAGG 2619  
1504 gcaacaggtccaggatggtgcgtcttaacaaaatgcaactaacatcaggaacttcaagt 1563  
2620 GCTATAGTTTCCAGGCTTGTGATTAATCCAGAAATATGCAAAATCTTGGGACAGGATA 2679  
1564 atcattccttataatagatggaagatgacatccctgaattcgaatattgctccagta 1623  
2680 ATTTATCTTTCATCTATGTTGGCACTAACACTGTTACTCTTAGCAATTTGTACCCATC 2739  
1624 ctgacgtgaagaatgattgaaccgcagcaatgactgtgattgccacaagaataag 1683  
2740 ATTGCAATAGCAGAGATGTTGAATGAATAATGTTGTTCTGGACAAAGCATGAAAGATAG 2799  
1684 caagaacttaagcagctgtggaagaatagcaactgaagccttggagaatatacgtactata 1743  
2800 AAGAAGCTAGAAAGGTGCTGGGAAGATGCTACTGACATATGATGATGATGATGATGATG 2859  
1744 gctgcattacaagaaggaagaagccttcgaagaatglatgaagaagatgctcagactcaa 1803  
2860 GTTCTTTGACTCAGAGCAAGAAATTTGACATATGATGATGATGATGATGATGATGATG 2919  
1804 cacaagaatattcgcgaagaagaacagatattggaagcgtttagcatltaagcaagc 1863  
2920 TACGAAACCTTTGAGAAAGACACACATCTTGTGAATTAATTTTCTTACCCAGGCA 2979  
1864 ttataatttgcctatgacagcaggtttgcattttagagacttaattcaagctgga 1923  
2980 ATGATGATTTTTCCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3039  
1924 cgaatgaccccaagagggcatgtacatagtttactcgaattgacatagatgagactagcc 1983  
3040 CTCATGACCTTTGAGAGATGTTCTGTTAGTATTTTACGCTGTTGCTGTTGCTGCTGCT 3099  
1984 atcggaagaagcgtcgtttagctccgaatataccaaagccaatcggggctgcgcacat 2043  
3100 GTGGGGCAAGTCAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3159  
2044 ctgttgcctgttggaaagaagaacaaatataagacaagcgcagtcgaagaaggaagaaag 2103  
3160 ATCATCATGATCATTTGAAGAAACCCCTTTGATGACACTTACAGCAGCGAAGGCTTATAG 3219  
2104 ccagacacatgltgaagggaatttagagtttcgagaagctcttcttcatccatgctgc 2163  
3220 CCGAACAATATGGAAGAAATGTCATATTTGATGATGATGATGATGATGATGATGATGATG 3279  
2164 ccagatgttcatcctccgtggtcttccctcagtagatgagcggaaggaagaacagtagca 2223  
3280 CCGGACATCCAGGCTCTTCAAGGACTGAGCTGAGGAGGAAGGAGGAGGAGGAGGAGGAGG 3339  
2224 ttgtgggagcagcagcgtgtgtgggaagaagcacttcgttcaactctcgaagacttat 2283  
3340 CTGATGGGACAGAGTGGCTGTGGAGAGCAACAGTGGTTCAGCTCTCTGAGCGGTTCTAC 3399  
2284 gaccgcgttgcagaagcagctgtttagatgtgtgtagtgcagaagaatltgaatgtagag 2343  
3400 GACCCCTTGGCAGGGAAGTGTGCTTATGAGCAAGAAATTAACGACGATGAAATGTTACG 3459  
2344 tggctcgtttcccaatagcaatcgttcccaagaagcgtgtgcttcaactgcaagcatt 2403  
3460 TGGCTCCGACACACCTGGGCTGCTGCTGCCAGAGCCCATCTGTTGACGACGACATTT 3519  
2404 gctggaacatcgctatggtgtagaacaagcgtgtgtgtgccaatagatgagatcaagaag 2463

```

|||||
Db 3520 GCTGAGACATTGCTATGAGACAACAGCCGGTGCTGCACAGGAAGATTGTGAGG 3579
QY 2464 ggcgcaaatgcaagcaaatatccattcttattgaaggtccctcgagaataacacaca 2523
Db 3580 GGAGCAAAAGGAGGACACATACATGCTTCATGAGTCACTGCTTAATAATATAGCAGT 3639
QY 2524 caagttgacgtgaaggaacacagcttctgctgagccagaacaaagactatgata 2583
Db 3640 AAAGTAGGAGACAAAGAACTAGCTCTTGTGGCCAGAAACACAGCTTGCCATAGCT 3699
QY 2584 agggctctctccaaaaccaaatttattgttggatgaggaagccactcagccctcgat 2643
Db 3700 CGTGCCCTGTTAGACAGCCTCATATTTTGTGGATTGAGAACCCAGCTGCTGAGT 3759
QY 2644 aatgacagtgaagagtggttcaagcatcccttgataaaagcagaagcggaagacatgc 2703
Db 3760 ACAGAAAGTGAAGAAAGTTGTCACAGAAAGCCCTGACAAAGCCAGAAAGCCGACCTGC 3819
QY 2704 ctagtgtcactcacaggtctctcgaattcagaagcagattgatagtgttctgcac 2763
Db 3820 ATTGTGATTGCTCACCGCTGTCACACATCCAGAAATGCACTTAATAGTGTTCAG 3879
QY 2764 aatgaaagataaagaaacaaagaaactcaagaagctcctgagaatcgagacataat 2823
Db 3880 AATGGCAGAGTCAAGAGAGATGGACAGCATCAGACGCTGTCGACACAGAAAGCATCTAT 3939
QY 2824 tttaagttagtgaatgcacagtcag 2848
Db 3940 TTTTCATGTGTCAGTGTCCAGCGCTG 3964

```

# RESULT 7

US-08-583-276-18

Sequence 18, Application US/08583276

Patent No. 5837536

GENERAL INFORMATION:

APPLICANT: McDonaldh, Kevin T.

APPLICANT: Nienhuis, Arthur

APPLICANT: Tolstoshev, Paul

TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN

TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED

TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES

NUMBER OF INVENTIONS: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carella, Byrne, Bain, Gillfillan,

ADDRESSEE: Cecchi & Stewart

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: DM4.V2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/583,276

FILING DATE: 05-JAN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/332,444

FILING DATE: 31-OCT-1994

APPLICATION NUMBER: 07/887,712

FILING DATE: 22-MAY-1992

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 4669 bases

TYPE: nucleic acid

STRANDEDNESS: singular

TOPOLOGY: linear

MOLECULE TYPE:

DESCRIPTION: Genomic DNA

US-08-583-276-18

Query Match 35.9%; Score 1025.4; DB 2; Length 4669;

Best Local Similarity 62.1%; Pred. No. 3.4e-285;

Matches 1692; Conservative 0; Mismatches 991; Indels 42; Gaps 3;

```

QY 166 attattgataaagaaccgtagataactttccacagctcgagataataactgatacc 225
Db 1526 ATTAATTGATTAATAAGCAAGTATGACAGCTATTCGAAGAAGTGCGCAAAACCAATAT 1585
QY 226 atagaaggaacgttggaaatlaaaaatgttcttcaattatccatcaagaccatcacc 285
Db 1586 ATTAAGGAAATTTGGAATTCAGAAATGTTCACTTGAATGATTCACCACTTCGAAAAAGATT 1645
QY 286 aagattctgaaagtcgaatcagaatlaaagtcctggaagacagtcgcttgctgctg 345
Db 1646 AAGATCTTGAAGGCGCTGAAGCTGAAGGTGACAGAGTGGGACAGCGGTCCCGCTGTTGGA 1705
QY 346 ctcaatgcaagtcggaagagtaagtcagtcagcttcgcaaggttatatgacccgat 405
Db 1706 AACAGTGGCTGGGGAAGACACACAGTCCAGCTGATCAGAGGCTATGACACCCACA 1765
QY 406 gatgcttatacatggttgatgagaaagaaatcaagagctttaaattgtgagcatatga 465
Db 1766 GAGGGATGCTCAGTGTTCATGAGACAGATTTAGACCATTAATATTAAGTTTCTACGG 1825
QY 466 gaccaattgagtcgttgaagagcctgttctggaagcacaacacagtaacaaat 525
Db 1826 GAAATCATTTGGTGTGTGATGACAGAACTGATTTGTCACACAGATAGCTGAAGAAC 1885
QY 526 atcaagatgacagagatgactgactgataagagaatgagagagcagaaggaagca 585
Db 1886 ATTGCGTATGCGCGCTGAAGAAATGACCATGAGAGATTTGGAAGAACTGTCAAGGAAGCC 1945
QY 586 aatgagatgattatcatctgagttctcctaataaatttaacatttgtagggagaanaa 645
Db 1946 AATGCTATGACTTTATTCATGAAACTGCTCATTAATTTGACACCTGTGTGAGAGAGA 2005
QY 646 ggaagcacaatgagtcgagagagcagaacagagatgacatgctgctgcttgaatgca 705
Db 2006 GGGGCCCATTTAGTGTGTGGGACAGAGAGATGCGCATTCAGAGTGGCTGCTGTTCCG 2065
QY 706 aacccaagatctgactttagatgagtgactgactgctgctgacttcaagaagaatca 765
Db 2066 AACCCCAAGATCCTCTGCTGTGATGAGGACACCTTGAGACACAGAAACCAAGACA 2125
QY 766 gctgttcaagctgcaactgagaaagcagaagagtcgagactacaatcgtgtgtagacac 825
Db 2126 GTGGTTTCAGGTGGCTTGGATTAAGGCCAAGAAAGGTCCGACACCATTTGTATAGCTCAT 2185
QY 826 cgaacttctactaattcgaagtcgagattgattgtgacccctaaaggaatggaatgcgag 885
Db 2186 CGTTTGTCTACAGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2245
QY 886 gagaagagagcagatgctgaactaatgcaaaacagagtcctatattcaactgctgag 945
Db 2246 GAGAAAGGAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2305
QY 946 tcaagagatatlaaaaaagctgagta-----cagaatgagtcgaatgacata 992
Db 2306 ATGCACACACAGCAAGAAATGAAGTTAGAAATGACAGCTGATGATGATGATGATGATG 2365
QY 993 ttctactgaagaagaaccacacacacacacacacacacacacacacacacacacacacac 1052
Db 2366 ATTGATGCTTGTGAATATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATG 2425
QY 1053 ctctactgacaagagtcgaggaatcagcc-----aatct 1086
Db 2426 ACTCGTAGAGAGTGTCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2485
QY 1087 aaagagataagtcctcctgagtcctcctataataaatttaagttaaacaagcctgaa 1146

```

Db 2486 CTGGATGAAGATACCTCCAGTTTCCTTTGAGAGATTAATGAAAGCTTAATTTAACTGAA 2545  
1147 tggccttcttggtctcgggacatbgtcttcgtcttaataatgaactgtctccagta 1206  
2546 TGGCCTATTATTGTGTGTGTATTTTGTGCAATTAATATGAGAGCCCTGCACACGCA 2605  
1207 ttctcatcactcttgcaaaaataataacacgtcttggaataataataaac--aca 1263  
2606 TTTGCATATATATTTTCAAGATTAATAGGGTTTTCAGAAATATGATGATCCGTAACA 2665  
1264 ttaagcatatgacagaattatcatatgatatctgcatcttggtgttatcttct 1323  
2666 AAACGACAGATACATACCTTTGCTTCTACATGTTTCTACCCCTTGAAATATTTCTTT 2725  
1324 gtaagttatctatgacaggtatatttaagcagacaggaagaaatttaacagtaga 1383  
2726 ATTACATTTTCTCTTCAGGGTTTCACATTGGCAAGCTGGAGAGATCCACCAAGCGG 2785  
1384 ttaagacacttggtcctcaaaagcatgtatatacagagatactgcttggttgtaaaag 1443  
2786 CTCCGATATACATGGTTTCCGATCATGCTCAGACAGATGAGATGTTGATGACCT 2845  
1444 gaaacagcagcagaggtctgacacaaatattagccatagatatagacaataatcaagga 1503  
2846 AAAAACACACTGGAGATTGACTACAGGCTCGCCATATAGCTGCTCAAGTTTAAAGGG 2905  
1504 gcaacaggtctcagagatctgctcttaacaaaatgcaactaacaagagactctcagt 1563  
2906 GCTATAGGTTCCAGGCTTGTGCTATATACCAAGAAATACCAATTTCTGGACAGGATA 2965  
1564 atcaattccttataatagatgagagatgacattcctgattctgaglatgtctcagta 1623  
2966 ATTATATCTTCACTTGTGTGCACTAACTGTTACTTACGAAATTTGATCCCATC 3025  
1624 ctggcgtgacaggaatgattgaaacgcagcaatgactgatttgcacaagaagataag 1683  
3026 ATTGCATATACAGAGATTGTTGAATGAAATGTTGTGTGACAGCACTGAAAGATPAG 3085  
1684 caagaacttaagcactgtagaagatagcaactgaactgtgagaaatatacgtcata 1743  
3086 AAGAACTAAGAGTGTGGAAGATGCTGATCTAGTACATAGAACAACTTCCGAACGCTT 3145  
1744 gctgcatatacaaggaagaaagccttcagagcaaatgtaagaagagatgctcagactaa 1803  
3146 GTTTCCTTGTACTGAGCAGCAAGATTGAAATATGATATGATGCTCAGAGTTTCCAGTACA 3205  
1804 caagaataatccctcgaagaagacacagatatattggaagcgttatgattcattcagcagtc 1863  
3206 TACAGAAATCTTTGAGGAAAGCACACATCTTTGAAATACATTTTCCCTCAGCAGGCA 3265  
1864 ttatcatcttctgctatgcaagcaggtttcagatttgagagcatttaattcaagctgga 1923  
3266 ATGATGATATTTTCTATGCTGATGTTTCCGTTTGAGAGCTACTTGTGCGCATATAA 3325  
1924 cgaatgaccccaagaggaatgattcattctcgaattgacatatgagagctagcgc 1983  
3326 CTATAGAGCTTGTGAGATGTTCTGTGTATTTTACCTGTTCTTGTGGTGCATGCGC 3385  
1984 atcggaataacgcctgcttggcctcctgaatatccaaagccaatccggggcctgcgat 2043  
3386 GTGGGGCAAGTACTTATTTGCTCTGACATATGCCAAAGCCAAATATATGACAGCCAC 3445  
2044 ctgcttgccttgtagaagaagaaacaatatagacagcccgatcaagaaggaagaaag 2103  
3446 ATCATTCATGATCATGAAAAACCCCTTTGATTTACAGCTACACAGCGAAGGCTATAG 3505  
2104 ccagacacatgtagaaggaatttagagtttcgagaagtcctctctctatccatctgc 2163  
3506 CCGAAACATTTGGAAGAAATGTACATTTGTGTAAGTGTATCACTACTCCACCGCA 3565  
2164 ccagatgttttcatcctccgtgctctcctcagtatgagcgaagaaagacagtagca 2223

Db 3566 CCGACATCCCATCTTCAAGGACACTGAGCTGAGGTGAAGAAAGGCCAGAGCTGGCT 3625  
2224 ttgtggagagcagcggctgttggaagaagcactctgttcaactctgcagagacttat 2283  
3626 CTGTGGGACAGAGTGTGCTGTGGAGACAGATGTGTCCAGCTCGTAGCGGTTTAC 3685  
2284 gaeccggtcgaaggaagatgctgttctgtatgtagtgaatgcaaaagatgtagtag 2343  
3686 GACCCCTTGGCAGGGAAGTGTGTGATGACAAATAATTAAGCAGCATGTTTCAG 3745  
2344 tggctcgttcccaatagatcgtctcgaagagcctgtgctctcaactcgaagcatt 2403  
3746 TGGCTCCGAGACACTTGGGATCTGTCCAGAGAGCCATCTCTTTGACTGACGATTT 3805  
2404 gctgagacatcgcctatggtgtaacaagcgtgtgtgtgcatatagatgaagaa 2463  
3806 GCTGAGAACATTTGCTTGTGACACACAGCGGGTGTGTACAGAAAGATCGTAGG 3865  
2464 gcgcgaatgcaagaatlatccatcttcttattgaaagctcctcgtgagaatacaaca 2523  
3866 GCAGCAAGAGAGGCCAATACATGCTTTCATCGAGTCACTGCTTAATAATATAGACT 3925  
2524 caattgactgaaagagacagactctctggtggcagagaacaagaactagctatgca 2583  
3926 AAGTAGAGACACAAAGAACTCACTCTGTGTGTGCGCAAGAACACATGCTATGACT 3985  
2584 aggtctctctccaaaaccaaaatttattgttgatgagagcactcagccctgat 2643  
3986 CGTCCCTTGTATACAGCCTCATATTTGCTTGTGATGAGCAGCAGCTCTGAT 4045  
2644 aatgaagatgagaagtggttgcagcagtccttgatataagccaggaaggaagacatgc 2703  
4046 ACAGAAAGTGAAGAGTTGTCCAAAGAGCCCTGACCAAAACCAAGAGGCCGACCTGC 4105  
2704 ctatgttactcagagctcctcgaatctgcaatcagaagcagatgtagatggtctgcac 2763  
4106 ATTGTATGTCTCACCGCTCTCCATCCATCCAGATGACAGACTTAATATGATGTTTCAG 4165  
2764 aatgaaagataaagaagaagaactcaatcaatgaagctcctgtagaaatcgagacatat 2823  
4166 AATGCAAGAGTCAAGCATGAGCATGACGATCAGCATGCTGTGCACAGAAAGGATCTAT 4225  
2824 tttaagttagtgaatgcagatcag 2848  
4226 TTTTCAATGTGCTGCTGCTCAGCCTG 4250

RESULT 8  
US-08-752-447-1  
: Sequence 1, Application US/08752447  
: Patent No. 5994088  
: GENERAL INFORMATION:  
: APPLICANT: Mechaner, Eugene  
: TITLE OF INVENTION: Methods and Reagents for Preparing and  
: USING IMMUNOLOGICAL AGENTS SPECIFIC FOR P-GLYCOPROTEIN  
: NUMBER OF SEQUENCES: 2  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: McDonnell Boenken Hulbert & Berghoff Ltd.  
: STREET: 300 South Wacker Drive, Seventh Floor  
: CITY: Chicago  
: STATE: Illinois  
: COUNTRY: USA  
: ZIP: 60606  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: PatentIn Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/752,447  
: FILING DATE: 15-NOV-1996  
: CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5994088nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 95,1121  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-9808  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4669 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..424  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4264  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 4265..4669  
US-08-752-447-1

Query Match 35.8%; Score 1022.2; DB 2; Length 4669;  
Best Local Similarity 62.0%; Pred. No. 2.9e-284;  
Matches 1690; Conservative 0; Mismatches 993; Indels 42; Gaps 3;

QY 166 attatttgtaaaaccccgatagataactttccacagctcgagatataatcgc 225  
DB 1526 ATTAATTGATTAATAAGCCAAAGATTGACACCTTTCCAAAGAGTGGGACAAACCAGTAAT 1585  
QY 226 atagaagagactgtggaattaaatgttcttcaatcacaagaacatcaltc 285  
DB 1586 ATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCACATCTCGAAAGAAAGTT 1645  
QY 286 aagattctgaaagcttgatctcagaatlaagctcgagagaacagtcgctgtcgt 345  
DB 1646 AAGATCTTGAAAGGCGCTGAACCTGAAGGTGCAGAGTGGGACAGCGTGGCCCTGGTTGGA 1705  
QY 346 ctcaatgtcaggtggaagagtagtcgtagtcagagcttcgagagrtatagatcgcgat 405  
DB 1706 AACAGTGGCTGTGGGAAAGCACACAGTCCAGCTGATCAGAGGCTCTATGACCCACA 1765  
QY 406 gatgcttatacatgtgtgagatgagaatcagacagagcttlaaatgtgcgcatatcga 465  
DB 1766 GAGGGAGTGCAGTGTGATGAGACGATATTAGACCATTAATGTAAAGTTTCTACGG 1825  
QY 466 gacatatctgagatgtgtagtcaagaagcgtgttctgctggagcaacatcaaglaacaat 525  
DB 1826 GAATCATTTGGTGTGGTGAAGCAAGAACCTGTATTGTTGCCACGATAGCTGAAAC 1885  
QY 526 atcaagatgagcagagatgtagttagatgaaagatgagagagcaagaagggaaga 585  
DB 1886 ATTCGCTATGCGCGTGAATAATGTCCACATGATGAGATTGAGAAACCTGTCAAGGAAGC 1945  
QY 586 aatgctatgatttatactgagagttccctaaataatlaatacatctgtaggggaaaaa 645  
DB 1946 AATGCTATGACTTATCATATGAACCTGCTCAATAATTGGACACCTGGTTGGAGAGAA 2005  
QY 646 ggaagctcaaatgagtggaagagcaagaacagagatgcgaattgctgtgcttagttcga 705  
DB 2006 GGGGCGCAATTGAGTGGGAGAGAGAGAGATGCGCATTTGCAAGTCCCTGGTTCCG 2065  
QY 706 aaacccaagatctgatttagatgaggtctacgctcgcctgattcagaagaacagttta 765  
DB 2066 AACCCCAAGATCTCTGCTGATGATGAGCCACAGCTCAAGCTTGACACAGAAAGCCAGAA 2125  
QY 766 gctgttcaagctgcagctggaagagcgagcaaaagtcgagactacaatcgttgtagagcac 825  
DB 2126 GTGGTTACAGTGGCTGTGATTAAGGCCAAGAAAGTGGAGCACCATTTGTGATAGCTCAT 2185

QY 826 cgacttctactattcgaaagtgcagatttgattgtgacccctaaaggatggaatgctgcgcg 885  
DB 2186 CGTTTGTCTACAGTGTGTATGCTGACGTGACGTGCTGTTGCATATGAGAGCATTTGTG 2245  
QY 886 ggaagaagagacatgctgaactaagtcgcaaaacagaggtcatatcatcactgtgtag 945  
DB 2246 GAGAAAGGAATCATATGATGACCTCATGAAAGAAAGGCAATTACTTCAAACTTGTACACA 2305  
QY 946 tcaagagataataaaaaagctgtagacagatgtaglcaatgacatatcta----- 997  
DB 2306 ATGCACACAGCAGAGAAATGAATTAAGAAATGACAGCTGATGAATCCAAAGTGAA 2365  
QY 998 -----ctgaagaagaagccaactcactcctcctcgtcactctgtggaagagatca--- 1045  
DB 2366 ATTGATGCTTGGAATGCTTCATTAATGATTCAGATCCAGTCAATTAAGAAAAAGATCA 2425  
QY 1046 -----agtcagactcattgacaagctgaggaatccaccacaatct 1086  
DB 2426 ACTCGTAGAGTGTCCGTGATTCACAGCCCAACACAGAAAGCTTAGTACCAAGAGGCT 2485  
QY 1087 aaagagataagcttctcgtgaagctctcctataaaaaattlaaagtlaaacaagcctgaa 1146  
DB 2486 CTGAGTAGAAGATTAACCTCCAGTTCTCTTTGGAGAGATTATGAAGCTTAATTTAACTGAA 2545  
QY 1147 tggccttctgtgcttctggtggaacatgtgctcgtctcaatgaaatgtaactcgaatc 1206  
DB 2546 TGCCCTTAATTTTGTGTGTGTATTTGTGCATTAATTAAGAGCCCTGCAACAGCA 2605  
QY 1207 ttccatcatcctctgcaaaaattacaacatgttggaaatgataaaac-----aca 1263  
DB 2606 TTGCAATTAATTAATTTCAAAAGATTATAGGGTTTTCACAAAGATTGATGATCCTGAAAC 2665  
QY 1264 ttaaaagatgagcagaatattatccatgatatcgtcaatttgggttatttgcctt 1323  
DB 2666 AAACGACAGAAATAGTAATGTTTTCATATGTTTCTACCCCTTGGAATTAATTTCTTTT 2725  
QY 1324 gtcaattattcatgcaagagattatttcaagcagaagcagggaaatttlaagaagaga 1383  
DB 2726 ATTACATTTTCTCTCAAGGTTTTCACATTTGGCAAAAGCTGGAAGATCTCCACCAAGCG 2785  
QY 1384 ttaagacactgtgcttcaagaacatgatatcaagaatatgctcgtgttga tgaag 1443  
DB 2786 CTCCGATACATGTTTCCGATCCATGCTCAGACAGAGATGAGTTGTTTCAATGACCC 2845  
QY 1444 gaaaacagcagaagagcttgacaacaatatlaagcagatatagcaaatcacaagga 1503  
DB 2846 AAAACACCACTGAGCATTTGACTACAGGCTGCCAATGATGCTGCTCAAGTTAAAGGG 2905  
QY 1504 gcaacaggttccagagattggtcgttaacaacaagaatgaactaactatggaacttcaat 1563  
DB 2906 GCTATAGTTCAGGCTTGTGTAATTAACCAATAATATACCAATCTTGGGACAGGAATA 2965  
QY 1564 atcattccttataatgagatgagatgacatcttcgttctgtagtattgttccagta 1623  
DB 2966 ATTAATATCTCTTCAATGATGTTGGCAACTGATGACTGTACTTACCAATTTGATCCATC 3025  
QY 1624 ctgcccgtgacaaggaatgattgaaccgcagcaatgactggaatttgcacaagaataag 1683  
DB 3026 ATTGCAATGACAGAGATGTTGTAATGAATAATGTTGCTGACCAACACTGAAGAAATAG 3085  
QY 1684 caagaactcaagcatgctggaagaatagcaactcgaagagcttggagaataatagttactata 1743  
DB 3086 AAAGAACCTAGAAAGCTGTGGAGAGATCGCTACAGGAACCAATTAAGAACTTCCAAACGGT 3145  
QY 1744 gtgtcatlaacaagggaaaaagcctcgaagcaaatgatatgaagagatgcttcgagactcaa 1803  
DB 3146 GTTCTTTTGACTAGAGAGCAAGATTGGAACATATGATCTCAGAGTTTGCAAGGTACCA 3205  
QY 1804 cacagaataactcgaagaagcagagatbatggaagctgttataatgcatcaagcatgccc 1863  
DB 3206 TACAGAAACCTTTTGAGAAAGCACACATCTTTGGAATTAATTTCTTCTTCAACCCAGGCA 3265

OY	1864	tttataatttgcctatgcgaagggttcgtattggagccattttaatccaagctga	192
Db	3266	ATGATGTAATTTTTCCTATGCTGGAATGTTCCGGTTGGAGCCACTGGTGCCAACTAAA	3325
OY	1924	cgaaatgaccccaaggaggacatgttcataagtttttaactgcgaattgcatalatgagactagcc	1983
Db	3326	CTCATGAGCTTTGAGGATGTCTGTTAGTATTTTTCAGCTGTGTTGTCTTTGTCCTATGGCC	3385
OY	1984	atcggaaaaacgcctcglttttgctctcgtatataatccaagccaatcggggctgcgcatt	2043
Db	3386	GTGGGGCAAGTCAGTTTCATTGTGCTCCGATGATGCCAAAGCCAAAATATCACACACCAC	3445
OY	2044	ctgtttgccttggttggaagaagaaccaatacatagacagccgacccaagccaagaaggaagaag	2103
Db	3446	ATCATCATGATCATTTGAAAAAACCCCTTTTATTGATACAGCTACACGACGGAAAGCCTTAATG	3505
OY	2104	ccagacacatgttgaaaggaatttagagtctgcgaagctctcttccttcataccatgtgc	2163
Db	3506	CCGAACACATTTGAAAGAAATGTACATTTTGTGAAGTTGTATTCACATATCCACCCGA	3565
OY	2164	ccgaatgttttcaatccctccgtgtctatacccttaagtaattgagcgagaaagacagtagca	2222
Db	3566	CCGGAACATCCCAAGTCCTTCAGGGACGTGAGCCTGAGGTGAAGAAAGGCCAGACGGTTGCT	3625
OY	2224	ttcttgaggagcaagcgactgtggaaagaacacttcgttcaactcttgagaagactttat	2283
Db	3626	CTGTTGGGCACACATGGCTGTGGGAAGACACAGTGTGTCCAGCTCTGTGAAGCGGTTTAC	3685
OY	2284	gaaccocgtgcgaaggaacaagtgcgtgttbtatgltgtgtgatalgcgaagaagatttgaatgaag	2343
Db	3686	GACCCCTGTGGCAGGAAAGTGTGCTTGATGACCAAAATTAAGGAGCATGAATGTTGAC	3745
OY	2344	tgtgtccgtcttcccaataagaaatcgttccccaagaagccctgtctctcaactgcgaagcatt	2403
Db	3746	TGGCTCCGAGACACCTGGGCATCTGTGTCCAGAGAACCCATCCTGTTTACATGCGACATTT	3805
OY	2404	gctagagaacatcgcactatgttgacaacagccgtgtgtgtccattagatagataccaagaa	2463
Db	3806	GCTGGAACATTTGGCTATGTGAGAACACAGCGGGGTGTGTCAAGSAAAGATTCGTGAG	3865
OY	2464	gctgcgaatctgcgaacaatataatcattctttatttgaagttctcccttgagaataataacaca	2522
Db	3866	GCACAAAGGAGGCCAACAATACATATGCTTCATCGATCTGCTAATTAATATGACACT	3925
OY	2524	caagtttgaaactgaagaagagaacagcttctcggcgccaaacaaagaactatgactcttga	2585
Db	3926	AAATAGSAGACACAAAGAACTCAGCTCTCTGGGGCCGAAGAACAAAGCAATTCOCATAGCT	3985
OY	2584	agggctctctcccaaaaaaccaaaatttatacttltgtatgtaggcacactcaagccctgat	2643
Db	3986	CGTGCCCTTGTTAGACAGCCATATTTTTCCTTTTGGATGAAGCCAGCACGTACACTTGAT	4045
OY	2644	aatgaacagtgaagaagtgtgtctacagcatgcgccttgataaagccaagaacgggaagacatgc	2703
Db	4046	ACAAATAAGTGAATAAGGTTTGTTCACAAAGGCCCTTGACAAAGACGAGAAAGCGCCGACCTGC	4105
OY	2704	ctagttgttcaactcaagagctctctgcgaatttcgaagaacgaagagatttagttgtttcgac	2765
Db	4106	ATTGTGATTTGCTCACCGCCTGTCCACCATTCAGAAATGCGACTTAATTAATGTGTTTCAG	4165
OY	2764	aatggaagaataaagaacaagaagaaactcatcaagaagctctcttgaagaatcagagacatat	2823
Db	4166	AATGGCAGAGTCAAGGAGCAATGGACGACATCACACACTGCTGTGGCACAGAAAGCATTTAT	4222
OY	2824	tttaagttaagtgaatgcacagctcaag	2848
Db	4226	TTTTCAATGGTCAAGTGTCCAGGCTG	4250

RESULT 9  
US-09-120-513-1  
; Sequence 1, Application US/09120513  
; Patent No. 6025160

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	Similarity	Score	95.3 %	DB 3	Length	4233	
Best Local	Similarity	60.4 %	Pred.	No. 2e-264			
Matches	1048	Conservative	0	Mismatches	1043	Indels	39
							Gaps
							3
QY	166	attatgtataaagaaccagataagataaactttccacagcttgatataaactgaatcc	225				
Db	1121	ATAATTGATAATGAGCGCAAGCAFTTGACACTTCTCAACCAAGGACACACAAACAGCACT	1180				
QY	226	atagaaggaaactctggaatttaaaatgtttccttcaattatcatcaagaccatcatc	285				
Db	1181	ATAAAGGGGAAATTTGGAAATTTAAAAATGTTTCTTCACATACCACATCAAGAGTGAAATT	1240				
QY	286	aagatcttgaagaagcttgatctcgaatctcgaatctgaagaagacagctgccttggtcg	345				
Db	1241	AAGATCTTGAAGAGGCGCTCACTGAAAGTGAAAGACGGGCAACAGGTAGCCCTGTTGGC	1300				
QY	346	ctcaatgycagctggygaagaagtaacggtatgtccagcttcttgcagaagttatatatccgat	405				
Db	1301	AACAGTGGCTGTGGGAAAGACAACTGTCAGCTGCTCAGAGGGCTTACGACCCCATTA	1360				
QY	406	gattgacctatcatatgttgatgagatgagaatgcacagctttaaagtgcgcatatcga	465				
Db	1361	GAGGGCGAGCTAGATCAACGACGACAGCAATCAGAGCAACATTAATGTGGATCTGCGG	1420				
QY	466	gaccatatggaagtgttatacaagaacctgttttcttgcggacaccatcaagtaaacat	525				
Db	1421	GAAATCATTTGGGGTGTGTGATCAGGAAACCGCTGCTGTTGGCACCACGATTTGCCGAAMC	1480				
QY	526	atcaagatgbyacgaatgatgttgactgtatgaagaagtgaagaagcaagcaaggaagca	585				









Db	2621	gtcttcgtgtggaattatctgaataagaactgtgtctgtgtcgaagctctgaaggaaga	2600
Qy	1687	gaacttaagcaatctctggaagaagatagaacacttgaagctcttggagaataatatactgataatagt	1746
Db	2681	gagctagaagatctctctggaagatcgtctacagaagaattgaanaacttcgcgaactgtgttc	2740
Qy	1747	tcattaaacaagggaanaaagcctctcgaagcaaatgataagaagatgcttcgaactcaaac	1806
Db	2741	tccttcgactagggagacagaagaatttgganaatactgatactgcagagactcttcagataccatac	2800
Qy	1807	agaaataccccaagaagaagaacagataatttggaaagctgttatgcaattcaagcaagccttc	1866
Db	2801	agaaatgcttttgaagaagaagaacacgctcttttggatacaaccttcgccttcacccaagccagt	2860
Qy	1867	atacaatttgcctatbcagaagagtttcgcattcttgagagccttatataatccaagctcgagaca	1926
Db	2861	atttattttctcatagtctgtgtttccgcgttcgcgttcactacttggtcgaacgagaactc	2920
Qy	1927	atgaccccaaggagcagtctcatagttcatagtttctaactgcaattgcatatgagctatgagccatc	1986
Db	2921	atgacgcttttgaanaatgttaagtgtgtaatttctcgtctgtctgtcttcttggatcgagaca	2980
Qy	1987	ggaaaaacgctcgttttggctcttgataatttcacaagaacaaatcsggggcgtcgcatctg	2046
Db	2981	gggaataaccagttcatatctgcctcttcgactacgcgaaagccaaagttctcagatcccaatcc	3040
Qy	2047	tttcgcttcttggaaagaagaaccacaataatagaacgcgcgagttcaagagaagggaanaagcca	2106
Db	3041	atcagagttcatcttgagaanaatccccgcgatttgacagctacagaacgagaggtcttgaagctc	3100
Qy	2107	gacacatgtgaagggaatttaagattctcgagaagctctcttctcatocatalgtoccca	2166
Db	3101	aattgtttagaagaagaatgtgaatttaattgaagttcatgtcatatcccaaccgaagccc	3160
Qy	2167	gagttttcaactctccgttgctgtctatccctcagtatattggcgcgagggaaagacaagtagcattt	2226
Db	3161	aacatcccaagtcgtccaaggagctcgaagcttcctcaggtgaagaaggggcaaacgcttcgcctg	3220
Qy	2227	gtggagggagcagcgctctgtggaagaagcacttcgtttcaacattctcgaaagactttatgac	2286
Db	3221	gtgggcagcagctgtgctgtcgcgggaaagatacagtggtgcagctgtctcgagcgcttctataac	3280
Qy	2287	cccgctgcagaagacaagtgctgttcttgaatgtgtgtgaatgcgaagaagaatctgaattgaacgtg	2346
Db	3281	cccatctgtggaacaagtggttctctagaatgcgaagaagaataaacaactcaacgtccagctg	3340
Qy	2347	ctccggttcccaaatagaacatcgttctctcaagcgctgtgcttcttaactgtcagcatgtct	2406
Db	3341	ctccgcgcgccaactcgtggaacttctgttcccaagcgagcccatctcttctgtactgtgcagctaccc	3400
Qy	2407	gagagacatgcctatagtgtgacaacacgcccgtgtgtgtcccaattagatgagatccaagaagcc	2466
Db	3401	gagagaactctgcctacgcgagagacaacacgctgtctcgtgtctcatgagagagatcgttgaagctc	3460
Qy	2467	gcaaatgcagcaaatataccattcttcttatttgaagtgctctccctggaanaatacaacaacaaa	2526
Db	3461	gccaaggagagccaacatccacatccacgacttcatcgcactcactcgtcgtgagaaatacaacaaccca	3520
Qy	2527	gtctgactgaagagagacacagcttctctgtgcgcgcgaagaacaagaactgataatgtgacaag	2586
Db	3521	gttggagacaacaagagctcgcgtctgtcgggcgcggagaaagacagctacatgcctacgcgcgc	3580
Qy	2587	gctctctcccaaaaacccaataatttatatgttggatgagagccaactctgaagcctctgataat	2646
Db	3581	gacctctgcacagcagctccacaacttaacttctgtgataagcgcacatccacgctctgtgatacg	3640
Qy	2647	gacacgtgagaaggttggttctcgaatctgtcccttgaataaagccaagacagcgagaagagacatccta	2706
Db	3641	gagagtgaaaaagtgctgtccaggaagcgctgcggaacaagccagggaaagcgccgaacctgtgatt	3700
Qy	2707	gtgtgtcaactcaagcgtctctctgcaattcaatccaagcagcagattgtatagtgtgtcttcgacaat	2766
Db	3701	gtgtgtcgcgcagccgctgttccacaactcccaagaaacgcagacttbatctgtgtgtatccaagac	3760

Oy	2767	ggaagaagtaagagacacaggaactcatcaagagctctcttgagaaatcgagacatatattt	2826
Db	3761	ggccagagtcacagagacgcgcgcgcacacccacacagcagctctgcgccagaagaagcatctattc	3820
Oy	2827	aagtlagtgaaatgcacacagtcacgttcagttga	2856
Db	3821	tcgatatcttcagagcttgagacaaagcgcctca	3850

RESULT 11  
US-08-461-823-1

; Sequence 1, Application US/08461823  
; Patent No. 5593840

```

; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
;

```

APPLICANT: George Jr., Albert L  
; Nazarenko, Irina  
APPLICANT: Nazarenko, Irina  
; Nazarenko, Irina

; TITLE OF INVENTION: AMPLIFICATION  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OncorPharm, Inc.  
;; STREET: 300 Barry Parkway

SINCE: 200 Ferry Parkway  
CITY: Gaithersburg  
STATE: Maryland

CNTRY: USA  
ZIP: 20877

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;

```

```

;
;
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-D
;

```

```

; SOFTWARE: PatentIn Release #1
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/461  
FILING DATE: 05-JUN-1995

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/1668  
FILING DATE: 16-DEC-1993

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/010  
;; FILING DATE: 03 JAN 1993

FILED DATE: 2/-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Karta Clorn E

NAME: NALCA, GIENH E.  
REGISTRATION NUMBER: 30,649  
REFERENCE/DOCKET NUMBER: PA-0

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301 527-2058

TELEFAX: 301 208-6997  
INFORMATION FOR SEO ID NO: 1:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
;

```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

```
;      HYPOTHETICAL: NO  
;  
; ANTI-SENSE: NO
```

US-08-461-823-1

Query Match	30.08; Scc
-------------	------------

Best Local Similarity 61.7%; pre  
Matches 1438; Conservative 0;

QY 561 gatggagagcagcaaggygaagcaaa

Db 1 GATGAGAAAGCTGTCAAGGAAGCCAA

Qy 621 attaatatctgttaggggaaaagg

Db 61 ATTTGACACCCCTGGTTGGAGAGAGAGG

```

Oy 2767 ggaagataaaggaacaaagaaactatcaagaagctcctgagaatacgcgaacatatttt 2826
      || || || || || || || || || || || || || || || || || || || || || ||
Db 3761 ggcagagcaagaagacagcagccaccaccagcagctcgcccaagaagacatatttc 3820
      || || || || || || || || || || || || || || || || || || || || || ||

Oy 2827 aagtagtgtaagtcacagtcagtgacgga 2856
      | || | | | | | | | | | | | | | | | |
Db 3821 tcaatggttcagctgagcagaagacgctca 3850
      | || | | | | | | | | | | | | | | | |

RESULT 11
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,823
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/168,621
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,433
; FILING DATE: 27-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Karta, Glenn E.
; REGISTRATION NUMBER: 30,649
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301 527-2058
; TELEFAX: 301 208-6997
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-461-823-1

Query Match 30.0%; Score 856; DB 1; Length 2726;
Best Local Similarity 61.7%; Pred. No. 1,6e-236;
Matches 1438; Conservative 0; Mismatches 850; Indels 42; Gaps 3;

Oy 561 gatggagagagcagcaagggaaagcaaatgctatgatttcatcagtggttcctaataa 620
      || || || || || || || || || || || || || || || || || || || || || ||
Db 1 GATTGAGAAAGCTGTCACAGGAAGCAATGCCATGACTTATCATGAAACTGCTCATTA 60

Oy 621 attcatatcatgtgtaggggaaagagagctcaaatgagtggagggcagaaacagaagat 680
      || || || || || || || || || || || || || || || || || || || || || ||
Db 61 ATTTCACACCCCTGTTGGAGAGAGGGGCCCACTTGAATGATGTGGGACAGAAACAGAGAT 120

```

OY	681	cgatattgcctgcgacctagtttgaaaccccaagatctcgaatttagatgagctacgctc	740
Db	121	CGCCATTGCAAGTCCCTGGTTGCAACCCCAAGATCTCTCGTGGTAGAGCCACGTC	180
OY	741	tgccttgatccaagaagcaagtcacgctgltcaagctgacgtggaagagcgacaag	800
Db	181	AGCCTTGACACACAAAGGACACAGTGGTTCAAGTGGCTTGGAATTAAGCCAGAAAGG	240
OY	801	tcgagatacaatcgtlgtlagaacacggaacttctactatcgaaatgcaattgattgt	860
Db	241	TGCGACACACCTTGTGATACCTCATGCTTTGCTACAGTTCTTAATGCTGAGCATGTCG	300
OY	861	gacctaaagatggaatgctgctgcggaagaagagacatgctgacataatgccaacg	920
Db	301	TGCTTTGATGATGAGATCTCATTTGGGAGAAAGAAATCATGATGAATCTATGAAGADA	360
OY	921	agctcatattatcacctgtgltgcacagagatatlaaaaagctgataccaagatg	980
Db	361	AGGCATTTACTTCAAACTTGTTCACATGCAAGACAGAGAAATGAAGTTGATTTGAAAA	420
OY	981	gtcaatgacatatctca-----ctgaagaagaaccaactacttcct	1023
Db	421	TGCGAGCTGATGAATCCAAAAGTGAATTTGATGCTTGGAAATGCTTCAAAATGATTCAAG	480
OY	1024	ctgacactctggaagagacatca-----agtaactattatga	1065
Db	481	ATCCAGTCTAATTAAGAAAAAGATCACTGTAGAGCTGCCGTGCATCAAGCCCAAGA	540
OY	1062	caagctgaggaatcaccccatcatcaagaagataagctctctcgaagctcctctataa	1121
Db	541	CAGAAAGCTTGTATGCCAAAGAGGCTCGATGAAATATACCTCCAGTTTCTTTGGAG	600
OY	1122	aatttaagltaaacaagcctgaatgacctgttgtgtctcgtggaacatgacctgt	1181
Db	601	GATTATGACGCTAATTTTAATCAATGAGCCTTAATTTGTTGGTGATTTTGTGCCAT	660
OY	1182	tctaaatggaactgttcatccagatatttccatcacacttggcaaaaatatatacagtg	1241
Db	661	TATAAATGAGGCGCTGCAACCGATTTTGGAAATTAATTTTCAAAAGATTAATGAGGGTTT	720
OY	1242	t---ggaatatgataaaacacatlaaagca tga tgcagaatattatccatgatat	1298
Db	721	TACAAGATTTGATGATCTCTGAAACCAAAACGACAGATAGTAACTGTTTTCATGTTGT	780
OY	1299	cgcaatttgggtgatttattgcttgtgcagtatctcaatgcagggatattatlcagcgag	1358
Db	781	TCTAGCCCTTGGAAATTAATTTCTTTTATACATTTTCTTCACAGGGTTTCCACATTTGGCAA	840
OY	1359	agcaggggaaaatttaacgagtatgaagacacctggcctccaagaagcattgatatca	1418
Db	841	AGCTGGAGAGATCTCTCCAAAGGGGCTCCGATTCATAGTGTTTTCCGATTCATCTTGACA	900
OY	1419	ggaattgcgcggttctgataaagaagaaacagcacaggaagctgtgacacatatlga	1478
Db	901	GGATGTAGTGGTTTATGATGACCTTAATAAACACACTGAGAGCTTATGACTACAGGCTGCG	960
OY	1479	catagatatagcacaatltcaaggagcaacaggttccagagatggcgtcttaacaaaa	1538
Db	961	CAATGATGCTCTCAATTTAAAGGGGCTATAGGTTCCAGGGCTGCTGTAAATTACCCAAA	1020
OY	1539	tgcacatacaatggaacttccgttatcatcttcccttatatatatgagtggatgaact	1599
Db	1021	TATACCAAAATTTGGGACAGAAATTAATTAATCTTCACTATAGTGTGGCAACTAACCT	1080
OY	1599	cctgattctgagtatgtctcagtaactgtgcgagcaaggaatgattgaaacccgagcaat	1655
Db	1081	GTTACTCTTAGCAATTTGATCCCATCTTGCATAATGACAGGGAGTGTGGAAATGAAATGTT	1140
OY	1659	gacttgatttccacaagaataagcaagaacttaagctatgctggaagaatagacaact	1718
Db	1141	GTCGTGCAACACCTGGAAGGTAAGAAGAACTAGAAAGGTGCTGGGAAATATGCTACTGTA	1200
OY	1719	agctttggaataatacgtactaagtgtcatlaaacaaggaaaaagccttcgagcaaat	1778

Db	1201	AGCATATGAAAACCTTCCGAACCGTTTCTTCTTGACTCAGAGCAGAAAGTTTGAACATAT	1260
Qy	1779	gtatgaagagaatgcttcttoagaactcaacacagaataactctgaagaagacacagatlatbgt	1838
Db	1261	GTATGCTCAGAGTTTGCAAGTACCATACAGAAACHTTTGAGAAAGAACACATCTTTGG	1320
Qy	1839	aagctgtlatgcatcgaatcgaatgaccttatatatcttgcctatgtgaagaaggttgatc	1898
Db	1321	AATTACATTTTCTTCACCCAGGCAATGATGATATTTTTCATAGCTGGATTTTCCGGTT	1380
Qy	1809	tggagaccttattaatcaagctgagcgaatgtgccccagaagggcagttcatcagttcttac	1958
Db	1381	TGGAGCTACTTGSTGGCACATATAACTCATGACCTTTGAGATGTTCTTGTTAGTATTTTC	1440
Qy	1959	tgcaattgcatalatgagcatatggccalcggaanaaacgctcglttbgctctgtaatlc	2018
Db	1441	AGCTGTTGCTTTTGTTGGTSCCATGSGCCGTGGGGCAAGCATGTTCAATTTGCTCGCTGACTATGC	1500
Qy	2019	caaaagccaatcggggggctbgcgcatcgttgccttglttggaanaagaaaccaatataga	2078
Db	1501	CAAAAGCCAAATATCACACACCACCATCATCATGATCATATGAAAAACCCCTTTGATTTGA	1560
Qy	2079	cagcgcgagtcacagaagaagggaanaagccagacacatbgtgaaggaaattagagtttcgaga	2138
Db	1561	CAGCTACAGCACGGAAAGCCTTATGCCGAACACATTTGGAAAGAAATGTCACATTTGGTGA	1620
Qy	2139	agctctctctctcatccatcgtcccaagatgtttcatcctcgttgcctalcctcctgaag	2198
Db	1621	AGTTGTTATTAACATATCCACCCGACCGGACATCCAGTGGCTTCAGGAGTACAGCCTTGA	1680
Qy	2199	tattgagcgaagaaagacagtagcatcttbtggygagcagcggctbgtgggaanaagcattc	2258
Db	1681	GGTGAACAAAGGCCACAGACCTGTGCTGTGGGACACCATGTGCTTGGGAAGACACAGT	1740
Qy	2259	tgttcaactctctgcaagacttatabaccgcgtgcaagaacagcgtctgttgaatgagt	2318
Db	1741	GGTCCAGCTCTGTGACGGGTTCTTACACACCCCTTTGGCAGGAAAGTGTCTTGATGGCAA	1800
Qy	2319	ggaatgcanaaagaattgaatgatacagctgctcgttcccaatagacatcgttctcctaaga	2378
Db	1801	AGAAATTAAGGAGACTGAAATGTTGACGTGGCTCCGAGCACACTGGGCATCGTGTCCACAGA	1860
Qy	2379	gctctgctcttcaactcgcacgatbgtcgtgaagaacatcgcttatbgtgacaacagcgtgt	2438
Db	1861	GCCCATCTCTGTTTGACGACGACATTTCTGGAACATTTGCTTGGAGAACACAGCCGGGT	1920
Qy	2439	ggtcgcattagatgagatccaagaagccgcaaatgagcaaatatccatcttcttatgga	2498
Db	1921	GGTGTCCACAGGAAGATCTGTGAGGGCAGCAAAAGGAGGCCAACATATGCTTTCATCGA	1980
Qy	2499	aggtctccctbgaagataatacacacacaagaattbgaactbgaagaagcagcttctbgycg	2558
Db	1981	GTCACGTGGCTAATTAATATATGACACTAAAGTAGAGACAAAGAACACTCAGCTCTGTGGG	2040
Qy	2559	ccagaacaagaagactagctatbgtcaagggctcttctccaaaaaccacaatttcatglt	2618
Db	2041	CCAGAAACAACGCAATTCGCCATACCTGCTGCCCTGTTTACGACGCCCTCATATTTTGCTTTT	2100
Qy	2619	ggaatgaagccactcagccctctgataataagatbgaagaagtggtcttagcagctcccttga	2678
Db	2101	GGATGAAGCCACGTACACTTGGATACAGAAATGTAAAGGTTGTTCAGAAAGCCCTTGA	2160
Qy	2679	taaaagcaagacgaggaagacatgcttagtbgtaacatcaagcgtctctbgaattcaqaa	2738
Db	2161	CAAAAGCCAGAAAGGCGGACACCTGCAATTTGATGTTGCTCACGCCCTGTCCACATCCAAAA	2220
Qy	2739	cgcagatttgatgtagtltctgcacaaatggaagaataaaaggaacaagaactcalcaaga	2798
Db	2221	TGCAGACTTAATAGTGTGTTTTCAGATGGCAAGTCAAGGAGCATGTGCACGCATCAGCA	2280
Qy	2799	gtctccatgagaatctgagacatatattttaagttagttgaatgacacgtcaag	2848

Db 2281 GCTGCTGGCACAGAAAGCATCTATTTCATAGTGTCAGTCCAGCGCTG 2330

RESULT 12  
US-08-996-545-1  
Sequence 1, Application US/08996545  
Patent No. 5928898  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atld of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,545  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4002  
US-08-996-545-1

Query Match 14.5%; Score 414.6; DB 2; Length 4002;  
Best Local Similarity 48.8%; Pred. No. 2.6e-109;  
Matches 1352; Conservative 0; Mismatches 1324; Indels 93; Gaps 5;

Db 1216 ATCGATCCCGAGTCCCATTCATATTCGAAAGGAAAGACCCCTGACCATTTT 1275

Qy 169 atcgatagaaccagatagataactttccacagctgatatataaactgataccta 228

Db 1216 ATCGATCCCGAGTCCCATTCATATTCGAAAGGAAAGACCCCTGACCATTTT 1275

Qy 229 gaagaaactgtggaatttaaaatgttcttcaattatcatcaagaacatctatcaag 288

Db 1216 GAGGCCACATTGAGTTACGCAATGTCAAGCATATTTACCCATCTAGACCGAGGTCAAC 1335

Qy 289 atctcgaaggtctgaaatcctagaatgaatctcgaagagacagctgccttgctgcctc 348

Db 1336 GTCAATGGAGATGTTCTCTGTCATATTCGCGTGGAAACAAACGCGTTTATGTCGGCCCC 1395

Qy 349 aatggcagtggaagagtagtagtgcagcttctgcagaggtatataatgcagatgat 408

Db 1396 TCTGCTCTGGAAAAAGTACGCTGCTGCTGTTGTTGAGCGATTCACATGCTGCTGCGC 1455

Qy 409 ggccttatcatcgttgatagaaatgacatcagagctttaaagtgcgcatattatcgagac 468

Db 1456 GGTACGGTTTGGTGGATGGCCATACATTCAGAACCTCATCTCCGCTGGCTTGCCAA 1515

Qy 469 cataatgagtagtgaatcgaagagctgttctgtcggagacacatcaatgaataat- 527

Db 1516 CAGATCTCTGTTGATGACGAGAGCGCTGCTGTTTGGACAGCATTTTAAAGAAATTT 1575

Qy 528 -----caagatagcagagatgtagtgcagatgaag 561

Db 1576 AGGCAGGCTTCATCGACGACAAAGTACGAGATATTCGAGATAGCTCCGGGAACTC 1635

Qy 562 atgagagagacgaaggaagaatcgatagatattatcatcagagttccataaa 621

Db 1636 ATCGAAGACGGGGCAAAATGCGCAATCTCATGACTTTATTAAGCTTGGCTTAAGT 1695

Qy 622 ttaatacatctgttagtgaagaaagagctcaaatgagtgagaggcagaaacagagatc 681

Db 1696 TATGAGACCAATGTTGGCGAGCGTGGCTTTCCTTCAGGTGGCCAGAAACAGCGCATTT 1755

Qy 682 gcaattgctgtgcttagtgcgaagaaaccccaagatctgattttagatgagctagctc 741

Db 1756 GCAATCGCCCGTGGCTTGTAGTACCCAAATTCCTCTGATGAAAGCTACTTCTG 1815

Qy 742 gccctgagatcagaagaagctcagctgttcaagctgacgtgagagagcgagaaagt 801

Db 1816 GCTTGGACACAAATCCGAAAGCGTGGTTCAAGAGCTTTGGACAGCGCAGCTTAAGC 1875

Qy 802 cggactaataatcgtttagcacaacagacttctactatctcgaagtgagatttgatg 861

Db 1876 CGAATCTATTTGATGATGCTGATGCTGCTTTCACAGATCAAAACGGCGACAAATTTGG 1935

Qy 862 accctaagatggaatgctgcgaggaagaagagacatgctgaactaagtcagaaaga 921

Db 1936 GTTCTGTCAATGCGAAATTTGCTGACAGCACTCAAGATTTGTTGACCGCGGA 1995

Qy 922 ggtctatattatcactgtgatgacagagatatataaa----- 963

Db 1996 GCGCTTATTCGCAATCTGTGAGGCTCAACGTATCAATGAACAGAGAGCTACGCC 2055

Qy 964 -----gctgatgaacagatgagatgacatattctactgaagaaag 1008

Db 2056 TTGGAGAGAGCGGACGCTGAGAGATCTCAAGATGAGATATTCGCAAAATCAAACTCGG 2115

Qy 1009 accaactcaactctctgcagactgtggaagacatcaagctcaagctcaatgacaagct 1068

Db 2116 TCAAGCGCATCATCCGATCTCGACGAAACCCACACCATTTAGCCGACGCCAC 2175

Qy 1069 gag---gaatccaccatcctaagaagatgctctcctgaagctctctatataaatt 1125

Db 2176 AAGTCTGTTTCCAGCGCATTTCTTAAAGACCCCGAAGCAACTCCGAATATCTCA 2235

Qy 1126 ttaag-----ttaaacagcctgaatgagccttttggt 1161

Db 2236 TTATGAGAGCGCTCAAAATTTGTTGCTCTTCAACCGCGCTGAATCCGATACATGCTC 2295

Qy 1162 ctggggaactggtctgtctcgaatgaatgaaactgttcatcagatatttccatcactt 1221

Db 2296 ATCGCTCTTGTCTTCTCAGTGTACTGCTGCTGACCAACCCACGCAAGCACTGCTAAT 2355

Qy 1222 gcaaaatataacat-----gttggaaataatgataaacacataaagcatg 1275

Db 2356 GCTTAAGCCATAGCACTCTCGCTCCAGATATACATATAGCAAGCTTGAACATGAT 2415

Qy 1276 gcaagaattatcatgatalctgcatlittggtgtatattgttgcagttatctc 1335

Db 2416 GCGGATTTCTGATCATGATGTTTTCGTTGGTGTGATCATTCAGTTATATCAGCAATCA 2475

Qy 1336 atgcaagattatttlaagcagagcagaggaatatttaagatgagatgaagacatg 1395

Db 2476 ACCAATGTTGCTGATTTGCGGTATGCTCCGAGAACTTATTTGCTGCGCGGAGAAAGCACT 2535

Qy 1396 gcttcaagacatgtatatacagatatgctgtgtttagatgaagaagaacagacaa 1455

Db 2536 GCCTTTCGACGATACCCCTCAAGACATTGCTTTTACACAGAGGAAGAAATAGCACC 2595  
Oy 1456 gggagccttgcaacaatattagacatatagatagacacaacttaagagagaaagctcttc 1515  
Db 2596 GGGGCTGTGACCTTCTTCCCTGTCACCGAGACGAGCATCTCTCCGGTGTAAAGCGGTG 2655  
Oy 1516 aggatggtgcttaacacaacttaacacttaacacttaacacttaacacttaacactta 1575  
Db 2656 ACCTAGACACGACATCTGATGACCTCCACGACCTTAGAGCGGGCTATCATTAATGCGCTG 2715  
Oy 1576 atatatgagagagagagacattcctgactgactgactgactgactgactgactgactg 1635  
Db 2716 GCGATTGCGGGAATGGGCTTATGTTGATCTCGGTTGCGGTTGCGGTTGCGGATGC 2775  
Oy 1636 ggaatgactgaacacgagacatgactgactgactgactgactgactgactgactgactg 1695  
Db 2776 GGTTCCTACGATTTCTATGCTATGCTAGCCAGTTTCAATCAGGCTCCAAAGCTTATGAG 2835  
Oy 1696 catgctgaaagatagacactgaagcttggagaataatactactactactactactacta 1755  
Db 2836 GGATCTGCAAACTTCTGCTGGAGGCTACATCGTCTATCGCACAGTTGGGTCTATTACC 2895  
Oy 1756 agggaaagacgcttcgagcaaatgtatgaaagagatgcttcagactcaacagaataacc 1815  
Db 2896 CCGGAAAGGAGTGTCTGGGAGATTACCATGCCAGCTTGACGACGAAGGAGGACACAGT 2955  
Oy 1816 tcgaagaagacacagattatggaagctgtatgactcagccactcactatatactttc 1875  
Db 2956 CTAACTCTGTCTGAGGTGATCCCTTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3015  
Oy 1876 gactatgacagaggttctgacttggagcacttaactaactgagcgaatgacccca 1935  
Db 3016 TGCCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3075  
Oy 1936 gaggagcgttcaactgacttactgactgactgactgactgactgactgactgactgact 1995  
Db 3076 TTCCGCTTCTTGTGTTTCTCTCGAGATCTCTTGTGCTCAATCGCGGGGACCCGTC 3135  
Oy 1996 ctgcttctgctcctgaatattccaagccaactcgggctgctgctgctgctgctgct 2055  
Db 3136 TTTTCTTGTGACGACGATGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3195  
Oy 2056 ttggaagaagaacaaatataagacagcgcgaagcaagaaggaagaagaagaagaacat 2115  
Db 3196 TTGACCCGAAAGCCCAAAATGATGATGCTGTAAGAGGCGGACGATCGAAAGCGTG 3255  
Oy 2116 gaaggagatttagagttcgagaagctctcttctctatcactgactgactgactgact 2175  
Db 3256 GAAGGTAAATCGAATTTAGGACGTCACCTCAGATACCCGACCCGCGGACGACGCT 3315  
Oy 2176 atctcgtgtgcttactcactgaatgagcgaagaagaagaagaagaagaagaagaaga 2235  
Db 3316 GTCTGCGCGGCTTGGACCTGACCGTGAAGCTGACAAATATGTTGGCTTGTGGACCC 3375  
Oy 2236 agcggctgtggaagaacacttctgtcaacttctgagaagactttagaccccggtcaa 2295  
Db 3376 AGCGGTGTGTGCAAGAGTACACCATTTGATGCTTACGCGCTTTTACGATGCGATTGCC 3435  
Oy 2296 ggaacaagt 2355  
Db 3436 GGGTCACTCTTGTGATGGGAAGACATTAAGTAACTAAATCAACCTCTTACCGGAGC 3495  
Oy 2356 caaatagcaactgcttctcactgaagacgctgtctcactgactgactgactgactgact 2415  
Db 3496 TTTTGTGCTAGTGTGACCGAGGACGACACTGTACCAAGGCGACCATCAAGGAAACATC 3555  
Oy 2416 gactatggtgacaacacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2475  
Db 3556 TTACTGTATTTGTGGAAGAGCTGACCGGAAGATTTCTGATTAAGGCTTGAAGGAC 3615  
Oy 2476 gcaaatatcacttcttattatgaagctctcctgagaataacacacacaaagtgtgactg 2535

Db 3616 GCTAATATCTACGACTTCATCATGCTGCCCGAGGGCTTTAATACATGTTTGGCAGC 3675  
Oy 2536 aagaagacagagcttctcgtgagcgaagaagaagaagaagaagaagaagaagaagaaga 2595  
Db 3676 AAGGAGGACATGTTGTGTGGCGGCAAAACGATGTGGCAATGTGGCCAGGCTTCTT 3735  
Oy 2596 caaaacccaataattattgt 2655  
Db 3736 CCGGATCCCAAAATCTCTTCTGATGAAGCGACGTCACCCCTCGACTCGGTCGAA 3795  
Oy 2656 aagtggttcaagctgactgacttgaagcgaagcgaagcgaagcgaagcgaagcgaag 2715  
Db 3796 AAGGTCTCCAGGCGGCTTGGATGCCCTGCCGAGGCGGACACATTCGCCCTTGCA 3855  
Oy 2716 cacaagctctgcaattcagaacgcaagattgtagtgttctgcacaatggaaagata 2775  
Db 3856 CACGCACTGACGACGATTTCAAAAGGCGGAGGATGATGATTTTTCGACCAAGCAAGATC 3915  
Oy 2776 aaggaacaagaactcactcagaagctctgagaatcgaacatgacatataatttaagt 2835  
Db 3916 GTCGAAAGCGGAACGACACGCACTGTCACGAAAGGCGGCTACTGAGGCTGCTC 3975  
Oy 2836 aatgcacag 2844  
Db 3976 AACTTGCA 3984

RESULT 13  
US-08-996-545-3  
Sequence 3, Application US/08996545  
Patent No. 592898  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,545  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-996-545-3

Query Match 14.5%; Score 414.6; DB 2; Length 4002;  
Best Local Similarity 35.8%; Pred. No. 2,6e-109;  
Matches 992; Conservative 360; Mismatches 1324; Indels 93; Gaps 5;

```
OY 169 atgataaagaaaccagatagataacttccacagcttgatataaactgaatcaca 228
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1216 AUCGACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 229 gaaggaactgtggaattataaagtcttccatcatcaagaacatcatcaaga 288
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1276 GAGGAGCCAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 289 attctgaagcttcgaactcagaattaaagctgagagagagctgctgtgtctc 348
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1336 GUCAGAGGAGGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 349 aatgagcaatggaagagtagctgtagcttcctgacagagttatcatgacgatgat 408
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1396 UCUGGUCUCUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 409 ggccttcatcgtgtgtagatgagaaatgacatcagaagcttaaatgtgagcatatcgagac 468
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1456 GGUAGCGGUUUGGUCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 469 catactgagtgtagttagtcaagagcctgtttgttgcgagaccatcatgaataaat- 527
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1516 CAGAUUCUUGUGUAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 528 -----caagatagacgagatgagtgtgactgataagag 561
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1576 AGCGACGGUCUCAGUCGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 562 atgagagagagcagaagagagagcaaatgtgatattatcatgagatttccataaa 621
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1636 AUCGAGAACGGCGGCAAAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 622 ttataatcgtgtgtaggagaaagagctcaaatgtagttaggagagagagagagc 681
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1696 UAUGAGACCAUAGUUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 682 gcaatgtctgtgccttagtgcgaataaccgaagattcgtatttagatgagctagct 741
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1756 GCAAUCCGCCGCGGCGGUUAGUGAGACCAAAAUUCCUGCCUGAGUAGAGAGAGAG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 742 ggcctggaattcagaagaagatcagctgttcaagctcgtcgtgagagagcgagaagct 801
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1816 GCCUUGGACACAAAUUCCGAGGCGGUGUUCAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 802 cggactacaatcgtgtgtagcaacacgacttctactatctggaagtgcagattgtgtg 861
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1876 CGAAGUCUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 862 accctaaagatggaatgtctgtagaagaagagacacatgctgaactaagtgcacaaga 921
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1936 GUUUCGUGCAUUGGCAAAAUUUGUGAGACAGAAUUCAGAUUUGUGUAGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 922 ggtctatatcttcaactgtgtatgtcacaagatataaaaa----- 963
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1996 GGGCGUUAUCCCAAAUUGUGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 964 -----gctgagaaagatgtagtcaatgacatattctactgaagaagaag 1008
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2056 UUGGAGGAGCGCGAGCGUGAGAGUUCAGAUUAGAUUAGGCAAAAUUAGAGAGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1009 acccaactactctctgtgactctgtggaagacatcaagtacagattcatgacaggt 1068
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2116 UCAACCGCAUUAUCCGAGUUCGAGGAGAAACCCAGCAUUGAGCGGAGCGGAGCCAC
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1069 gag---gaatcacccaatctaaagaagatagttctctgaagtctcttatataaat 1125
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2176 AAGUUCUUGUCCAGCGAGUUCUUAAGAAAGACCCCGGAAACAAUCCGAAAUACUCA 2235
```

```
OY 1126 ttaaag-----ttaacaagcctgaatggcctttgtgt 1161
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2236 UUAUGAGAGGUGUCUAAAUUUGUUGUCUUCUUCUUCUUCUUCUUCUUCUUCUUCU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1162 ctgggacattggtctctgttcttaaatggaactgttcatcagattttcatcatctt 1221
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2296 AUCGUCUUGUUCUUCAGUUGUAGUCUGUGUGGCGCAACCCAGCAGAGAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1222 gcaaaatataaccat-----gtttgaaataatgataaaccacattaaagatgat 1275
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2356 GCUAAAGCCAGCAGACUCCUGCCUCCAGAAUUCAGAUUAGAGAAACUUCGAGAU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1276 gcaaaatttaccatgatatattgctcatatttggtgtgatttgcctgaatttcc 1335
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2416 GCGAUUUCUGUUCAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1336 atcgagattattttaagcagagcagaggaatatttaagatgagatgaactgtg 1395
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2476 ACCAUUGGUGUGCAUUGGCGUAGUUGCGGAGAGACUUAUUGCGGCGAGAGCACU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1396 gcttcaaaagcagtattatcagagatatgtcgtgttggatgataaagaagaagcaca 1455
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2536 GCCUUCGAGAGUACUCCGUCAGAGACAUUGCUUUCUAGACAGAGAGAGAGUAGCAC
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1456 ggaagcttgacaacatatagaccatagatatagacaaattcaagagacaaagcttcc 1515
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2596 GCGCGUCGAGCCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCUUCU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1516 aggattgcttctaacaacaaatgcaaatcaatcagagacttccagttatcatcttct 1575
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2656 ACUCUAGGACAGAUUCUUAUGAGACUCCAGCAGACCCUAGAGAGGAGGAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1576 atatatgattgagagatgacatcttctgttctgtagattgtcccaagtcttgcgtgaca 1635
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2716 GCGAUUUGGAGAAAUUUGGCGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1636 ggaatgattgaacccgacagcaatgactgtatttgcacaagaagataagaactaag 1695
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2776 GGUUUCUACCGAUUCUUAUUGCUUACCCAGAGUUCUUAUACGCGUCCAGCUGCU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1696 catgttgaaagatagcaactggaacttcttgagaaatatacgtactactatgactaaca 1755
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2836 GGAUUCGCAACUUCUUGCGAGGAGCUUAGUUCUUAUCCGACAGUUGGCUUAGUAGC
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1756 agggaaaaagccttcgagcaaatgtatgaagaagatgcttcagactccaacagaataacc 1815
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2896 CGGGAAGGAGUAGUUGGAGAGAUUACCAUUGCCAGCUUAGAGCACAAGCGAGGAGCAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1816 tcgaagaagacacagattatttggaagcgtttagattcagcagcatgcttatattt 1875
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 2956 CUAAUUCUUGUUCUAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1876 gctatgacagaggttgcgaatttgcgaacttatttaattcaagctgagcaatgaccca 1935
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3016 UCGGUGGCGUUCGUGUUGUAGAGAGAGACACUUCUUGGACACAGAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1936 gaagagatgtcatagtttactcgaattgcatgtgagcgtatgagcgaatggaaaaag 1995
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3076 UUCGCUUUCUUGUUGUUCUCCGAGAUUUCUUCUUGGUGGCUCAUUCGCGGAGCAGC
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 1996 ctgcttttggtcctcgaatatccaaagcacaatcgagggtcgtgcatgtgttgcctg 2055
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3136 UUUUUCUUGGACAGACAGUAGUAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 2056 ttgaaaaagaaaccaatatagacagcgcaagtcagaagaaggaagaaagccagacatgt 2115
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3196 UUCGACCGAAAGCCACAAAUUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 2116 gaaggaattgagatttcgagaagatccttctctcatcatgtcgcgcagatgttctc 2175
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3256 GAAAGUGAAUUGAUUUAAGAAAGUGGACUUCAGAUUCCGACCCGCGGCAAGACGCU
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
OY 2176 atcctcgtgtgtattccctcagtatgtgagcagaagaacagtagacttgtggtgagc 2235
```

```

Db      3316   GUCUGCGGCGGCUUGGACCCUAGACCUGGAAGCCUGGACAUAUUGUGCGCUCUUUGCGAGCC 3375
              ||| : | | | | : | | : | | | | | | : | | : | | | |
Qy      2236   agcgcgtctgtagaaagcaactcttctgaacttcctgcagaagacttatgatgcccgctgcaa 2295
              ||||| : | | | | | | | | | | | | | | | | | | | | | |
Db      3376   AGCGGUGUGGGCAAGAGVUACCAACAUUGGACUUGGAGCGGUUUUAGAAGUCAUUGCC 3435
              ||| : | | | | | | | | | | | | | | | | | | | | | |
Qy      2296   ggaacagtctgcttttgatggtagtgcataaagaattgaatgtaacagtctgcctcc 2355
              || : | | : | | | | | | | | | | | | | | | | | | | |
Db      3436   GGCGUACUCCUUGUUNUGGGAAGAGACAAUAUAAAUAUAAUACUUCUACGCCAGC 3495
              || : | | : | | | | | | | | | | | | | | | | | | | |
Qy      2356   caaatagcaatcgcttccctcaagaagcctgtgctctccaactgcagcatgtctgtaaacatc 2415
              : | | : | | : | | | | | | | | | | | | | | | | | | |
Db      3496   UUUUCUGCACUGGUCACCCAGCGAGCCGACACUGUACCAAGGCACCAUUAAGAAAUVC 3555
              ||| : | | | | | | | | | | | | | | | | | | | | | |
Qy      2416   gccctatgttgacaacacagccgltgtygtccattagatlgatcataaagaagccgcaatgca 2475
              : | | : | | : | | | | | | | | | | | | | | | | | | |
Db      3556   UUACUUGUUAUUGUCGAAAGUAGCAGUACCGAAGAAUUCUUGAUUAAGCUCUUGCAAGAC 3615
              : | | : | | : | | | | | | | | | | | | | | | | | | |
Qy      2476   gcaaatccatctcttattatgtaaggctcccttgagaatatcacacacaagttgagcty 2535
              ||| : | | : | | : | | | | | | | | | | | | | | | | | |
Db      3616   GCUAUAUUCUACGACUUCANCANUUGUCUCCGGAAGCGCUYUUAUACAAGUUGUUGCAGC 3675
              ||| : | | : | | : | | | | | | | | | | | | | | | | | |
Qy      2536   aaagagacacagcttctgcgcgcgcgaacaaagaactgctaigtcaagagctctctc 2595
              || | | | : | | | | | | | | | | | | | | | | | | | |
Db      3676   AAGGAGGCAUUGUUGUGCGGCCCAAAGCAACGUGUGGCCAUUGGCCGAGCCCCUUCU 3735
              || | | | : | | | | | | | | | | | | | | | | | | | |
Qy      2596   caaaaaccacaatttatctgttgtgtagagccacttcagccctcgataatgacagtgag 2655
              || | | | | | | | | | | | | | | | | | | | | | | | |
Db      3736   CGGGAUCCCAAAUCCUUCUUCUGAUAGAACGACACUCCGCUUACUCCGAGUCAGAA 3795
              || | | | | | | | | | | | | | | | | | | | | | | | |
Qy      2656   aagtglyticagcatgcctctgatalaaagccgaagcgggaagacatgacctagtgtaact 2715
              || : | | : | | : | | : | | : | | : | | : | | : | | : |
Db      3796   AAGGUCGUCACGGCGGCUUGGAGUAGCGCTGCGCCGAGCCAACAAUCCGCCUUC 3855
              || : | | : | | : | | : | | : | | : | | : | | : | | : |
Qy      2716   caccagcctctgcgaattcgaagcagatlltagatgigtctgcacaatlgaagaata 2775
              || : | | : | | : | | : | | : | | : | | : | | : | | : |
Db      3856   CACCACACACACAGAUUCCAAAGCGGACGUGAUUCUUAUUGUUGACCAAGCAAGAVUC 3915
              || : | | : | | : | | : | | : | | : | | : | | : | | : |
Qy      2776   aaggaacaagaactcatlcaagagctccttgagaatacgcagacatabattaagtiagtq 2835
              || | | | | | | | | | | | | | | | | | | | | | | | |
Db      3916   GUCCAAACGGGAGACACAGAGAACUGUGUCCAGAAAAAGGCCGGUACUACGUGUC 3975
              || | | | | | | | | | | | | | | | | | | | | | | | |
Qy      2836   aatgcacag 2844
              || | | |
Db      3976   AACUUGCAG 3984
              || | | |

RESULT 14
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

Query Match	Best local similarity	Matches 1352; Conservative	14.5%; Score 414.6; DB 4; Length 4002;	48.8%; Pred. No. 2,6e-109;	0; Mismatches 1324; Indels 93; Gaps 5;
US-09-328-320-1					
APPLICATION NUMBER:	US/09/328,320				
FILING DATE:					
CLASSIFICATION:					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER:	08/996,545				
FILING DATE:					
ATTORNEY/AGENT INFORMATION:					
NAME:	Webster, Thomas D.				
REGISTRATION NUMBER:	39,872				
REFERENCE/DOCKET NUMBER:	X-11766				
TELECOMMUNICATION INFORMATION:					
TELEPHONE:	317-276-3334				
TELEFAX:	317-276-2763				
INFORMATION FOR SEQ ID NO:	1:				
SEQUENCE CHARACTERISTICS:					
LENGTH:	4002 base pairs				
TYPE:	nucleic acid				
STRANDEDNESS:	single				
TOPOLOGY:	linear				
MOLECULE TYPE:	CDNA				
HYPOTHETICAL:	NO				
ANTI-SENSE:	NO				
FEATURE:					
NAME/KEY:	CDS				
LOCATION:	1..4002				
US-09-328-320-1					
Query Match	14.5%; Score 414.6; DB 4; Length 4002;				
Best local similarity	48.8%; Pred. No. 2,6e-109;				
Matches 1352; Conservative	0; Mismatches 1324; Indels 93; Gaps 5;				
US-09-328-320-1					
169 attgataagaaccagtagatagataacatttccacagctgtagatataacctgaatccata	228				
1216 ATCGATCGCCATCCCATTAATGATCATATTGACAGAGGAGGAGCGCTGACCATTTT	1275				
229 gaagaacacgtggaatttaaaatforttccaattatccatcaagaacattatcaag	288				
1276 GAGGGCCACATTGAGTTACGCAATGTCAAGCATATTACCCTCTTAGACCCGAGGTCCAC	1335				
289 attctgaagagctgcatctcagaattaaagtcctggaagacagctgccttggtcgtc	348				
1336 GTCATGAGAGATGTTCTCTGTCATGCGCCGGTGGAAGAAACAACGCTTAGCGGCCCC	1395				
349 aatgacagtggaagagtagcagtagtcacagctctcgcagaggttatatgatccgatgat	408				
1396 TCTGCTCTGGAAGAAAGTACGGGTGCGCTTGTTGACCGATTCATCATGCTGTTGCG	1455				
409 gacttatcatgtgtagatagatagatcacaagagcttaaatcttgsggattatcagagc	468				
1456 GGTACGCTTTGCTGGATGCGCATGACATCMAAGACCTCAATCTCCGTGCTTCCGCA	1515				
469 catatcgagtagtgtagtcaagagcctgtttgttcgggaccacatcagtaacaalat-	527				
1516 CAGATCTCTTTGTTAGCCAGGAGCCTGTTCTTTTGGCAGACGATTTATAGAATATT	1575				
528 -----caegtatgacgagatgatgtgacgtatgaagsg	561				
1576 AGGCACAGGTCTCATCGGCACAAAGTACGAGAAATGATCCGAGATTAAGGTCCGGAACTC	1635				
562 atgagagagcagaagagaagaacaaatcgatatttatcatctgtagtlttccataaa	621				
1636 ATTCAGAACGCGCCAAATAATGCGGAATGCTCATGACTTTATTACTCCTTGCGTAGAGGT	1695				
622 ttaatacattgttaggggaaaaaagagctcaaatgagtagtgagggagcagaacagagatc	681				
1696 TATGAGACCAATTTGGGCAAGCTGGCTTTCTCTTTACAGGTGGCCAGAAACGCCATT	1755				
682 gcaattgctgtgcttcaattgtcgaaaaccccaagattctgattttagatgagagctacgtc	741				
1756 GCATAGCCCGGTCGCTGTTAGTGACCCCAAAATCCTGCTCTGATGAACCTACTTTCG	1815				
742 gccctgagattcagaagcaagtcagctggttcaagctgcaattgtagaagagcgagcaaggt	801				



Db 1816 GCGTTGGACACAAATCCGAGGCGGTTCAGACGCTTTGGAGAGGGCAGCTGAAGCC 1875  
QY 802 CGGactacaactcgtgtagcacacgcgaacttctactactcgaagtagcagatttgatg 861  
Db 1876 CCAACTACTATTGTGATTCGCTCATCGCTTTCACGATCAAAACGGCGCACAACTATGG 1935  
QY 862 accctaagaatgtagatgtagcgcggaagaagacacatgctgaactaattgcaaaag 921  
Db 1936 GTTCTGTGCTCAATGCAAAATTCCTGAAACAGAACTCACAGATGAATTTGGTACCGCGGA 1995  
QY 922 ggtcatalatctacacttgatgtagcacaggaattataaaaa----- 963  
Db 1996 GCGCGTTATCGGAACCTTTGGAGGCTCAACGATATCAATGAAACAGAAAGAGCTGACGCC 2055  
QY 964 -----gcgtatgaacagatgtagtcaatgacatatctactgaaagaag 1008  
Db 2056 TTGAGAGACGCGGACGCTGAGATCTCACGAATGCAATATTGCCAAATTCAAATTCGCG 2115  
QY 1009 accaactcaactctcctgtagcactctgaaagagacatcaatgactgaacttaagcaagct 1068  
Db 2116 TCAAGCGCATCATTCGATCTCGACGGAACCCACAAACATTGACCGCACGGCACCCAC 2175  
QY 1069 gag---gaatccaccacatcaagaagatgctctcgtgaagctcctcatataaaatt 1125  
Db 2176 AAGTGTGTTCCAGCGCGGATCTTTCTAAAGACCCCCCGAAACACTCGAAATATCTCA 2235  
QY 1126 ttaaa-----ttaacaacactgaatgagccttltggtt 1161  
Db 2236 TTATGAGACGCTGCTCAAAATTTGCTTCCTTCATCAACGCGCTGAATTCGATCATGCTC 2295  
QY 1162 cggggagacatggtctgttcttaaaatgaaactgttaccgatttccatcatctt 1221  
Db 2296 ATCGGCTTGTCTTCTTCACATGTTAGCTGGTGGGCCAACCCACGAGCAGCATATAT 2355  
QY 1222 gcaaaattatataccat-----gttggaaataatgataaaacacatlaaagacatg 1275  
Db 2356 GCTAAAGCCATACGACACTCTCCGCCAGAAATCATATAGCAAGCTTGACATGTAT 2415  
QY 1276 gcaagaattatcactgaatcgtcacttggtagtattgtctgtcagttatc 1335  
Db 2416 GCGGATTTCTGATGATGATGTTCTTCGTGGTGTATCATCTACGTTATCACGAGTCA 2475  
QY 1336 atgcagggattattttagcgacagagaggaattttaaagtagatgaagacactg 1395  
Db 2476 ACCAATGGTCTGATTTGCCGTATGCTCCGAGAGACTTATTCGTGCGCGAAGACACT 2535  
QY 1396 gacctcaagaacatgatatatacagatcattgcccgtgattgataaagaacacagaca 1455  
Db 2536 GCGTTTCCGACGATTAATCGGTCAAGACATTTGCTTTTGACAAAGAAAGATATACACC 2595  
QY 1456 ggaagctgacaacaatattagacatagataagacaanaattcaaggaacagatcc 1515  
Db 2596 GGGGCTCTACCTCTTCCCTGTCACGAGACGAACATCTCCCGTGTATACCGGTGTG 2655  
QY 1516 aggaattgggtctttaaacaanaaataagcaataatgagacttcaattacttccctt 1575  
Db 2656 ACTCTAGGACGATTTGTATGACCTCCACGACCCTAGAGAGGCGTATCATTTATTTGCCCTG 2715  
QY 1576 atatatagaatgagagatgacatcctgacttctgtagtctcgaacttgcgcgtgaca 1635  
Db 2716 GCGATTGGGTGGAAATTTGGCCTTAGTTTGTATCTCGGTGTGTGCGGTTCTCCGACATCG 2775  
QY 1636 ggaatgattgaaacgcgaacatgactgattgccaacaanaataagcaagaactaag 1695  
Db 2776 GGTTCCTACGATTTCTATGTGTAGCCAGTTTCATTCACGCTCCACGCTGCTTATGAG 2835  
QY 1696 catgctggagaagaatagcaactgaagcttggggaatatatcgtactatagtatataaa 1755  
Db 2836 GGAATCTGCAAACTTTCTGCGAGGCTACATGCTTATCCGCAAGTTGCGCTATTAAC 2895  
QY 1756 agggaaaaagccttcgagaaaatgtagaagatgcttcaagactcaacagaataacc 1815  
Db 2896 CCGGAAGGAGATGTCTGGAGATTTTACATGCCACGCTTGAGCACAAAGCAGACACT 2955  
QY 1816 tcgaagaagaacagatatttgaagcgtttagtaatacaacagcacttataattt 1875  
Db 2956 CTATCTCTGTCTTGGAGGTATCCCTGTTATATGCTCTCTGAGGCACTGTGTTCTTC 3015  
QY 1876 gctatgacgaaggttctgatttggagcctatttaactcaagctgaaagaccaca 1935  
Db 3016 TCGTTGCGCTCGGGTTTGTGGTACGGAGACACTCTTGTGTACCAAGATATGACATT 3075  
QY 1936 gagggcatgtcataagtttlaactgcaattgcatatgagactatgagcaatcgaaaag 1995  
Db 3076 TTCCGCTTCTTTGTTGTTTTCGCCAGATTCTTTGGTGTCAATCCGGGGCACCGTC 3135  
QY 1996 ctgcgtttagctccatgaataatccaaagcaaatcggggctgcgcactgtttccctg 2055  
Db 3136 TTTTCTTTGACACACATGAGGCAAGGCAAGAAAGGCGCGCCGCAATTCGACGACATG 3195  
QY 2056 ttggaaaagaacaaatagacagccgcagtcgaagaaggaaaagccagacacatgt 2115  
Db 3196 TTGACCCGAAAGCCACAAATTTGATTAATGCTGTGAAGAGGGCAGAACTCGAAACGCTG 3255  
QY 2116 gaaggaaattagaagtttcgagaagctcttcttctatccatgctgcagatgtttc 2175  
Db 3256 GAAGGTGAATGCAATTTTGAAGACGTGCACTTGATACCCGACCGCCAGAACAGCGCT 3315  
QY 2176 atctcgtgtgtatccctcaagatttagcgaagaagaacagtagcatltgtgggagc 2235  
Db 3316 GTCTTCGCGGCGCTTGACCTGACCGTGAAAGCCGTGCAATATGTTGCTGTGCGGACCC 3375  
QY 2236 agcggtctggggaagaagcctctgttcaactctgcaagactttagaaccgcgtgaca 2295  
Db 3376 AGCGGTTGTGGCAAGATACACATTCATTCATTCCTTGAGCGCTTTACGTTGCAATGCC 3435  
QY 2296 ggaacaagtgtgtttagatgtgtgtagtgcanaaagaattgaatgacagtgtccgtcc 2355  
Db 3436 GGGTCCATCTTGTGTATGGGAAGGACATTAAGTAATTAATTCATTCATTCATTCATTC 3495  
QY 2356 caaatgacaacgtcttcccaagaagcgtgtgtcttcaactgcaagatgtgtgaaatcc 2415  
Db 3496 TTTCTGTCACTGTGTGACGAGGAGCCGACACTGTACAGGGCCATCAAGGAAACATC 3555  
QY 2416 gctatgtagcaacaagccgtgtgtgtgctcatgtagagatacaagaacgcaaatgca 2475  
Db 3556 TTATCTTGATTTGTGCAAGATGACGTACCGGAAGAAATTTGATTAAGCTTGCAAGGAC 3615  
QY 2476 gcaaatatcatcttctttagaagttccctcgtgagaatatacaacaacaagtgtgact 2535  
Db 3616 GCTAATATCTACGACTTTCATCATGTCGCTCCGAGGGCTTTAATACAGTTGTGGCAGC 3675  
QY 2536 aaagagaacagcttctcgtgcggaagaacaaagactagctatgtcaaggtctcttc 2595  
Db 3676 AAGGGAGGATGTTGTCCTGCGGCGCAAAAGCAACGtGTGGCCATTTCCCGAGCCCTTCT 3735  
QY 2596 caaaaacccaaatttattgtttagtaggagcaacttccagccctgataatgacagtgag 2655  
Db 3736 CCGGATCCCAAAATCTTCTTCTCGATGAAGCACAGCTGACCTCCATCTCCAGTACAGA 3795  
QY 2656 aaggtgttcaagatgaccttgaataagccaggaacggaagacatgctagtgact 2715  
Db 3796 AAGGTGTCACAGCGGCTTTGATGCGCGTCCGAGGCGCAACCAATCCCGCTTGGCA 3855  
QY 2716 cacagactctctgcaatctcaagaacgagatgtatagtggttctgcacaatggaaagata 2775  
Db 3856 CACCGACTCAGACGATTTCAAAAGGCGGACGTTATCTATGTTTTCACCAAGCAAGATC 3915  
QY 2776 aaggacaaggaacatcatcaagaagcctcgtgagaacatcgagacatatatttaagtag 2835  
Db 3916 GTCGAAAGGGAAGCGCACGACGACTGTCCAGAAAAAGGCGGCTACTACAGCTGCTG 2895  
QY 2836 aatgcacag 2844  
Db 3976 AACTTGCAG 3984





Db 2536 GCCUUUCGAGCAUACUCCGUCAGACAUUUGCUUUUUAGACAGAAGAGAUAGCACC 2595  
Qy 1456 ggaagcttgacacaatatattagccatagataagaacaatccaaggagcaagttcc 1515  
Db 2556 GCGGCUUCGAGCCUUCUUCGCGCCACGAGAGAGCAUUCUUCGCGGUGUAGCGGUG 2655  
Qy 1516 aggatgtgcttctaacaacaattgacaatacatgtggacttcagttatcttcctt 1575  
Db 2656 ACUCUAGGACAGAUUUUAUGACUCCACAGACCCUAGAGCGGCUAUUUAUUGCCUG 2715  
Qy 1576 atataatgattggagatgacattccctgtattctgtattgtctcagttatcggtgaca 1635  
Db 2716 GCGAUUGGUGGAAUUGGCGCCUUAUUUGUACUUGCGUUGCCGUGUUCUGGCAUCC 2775  
Qy 1636 ggaatgattgaaacccgcgaatgtactgtattgtccacaacaagaataagcaagaacttaag 1695  
Db 2776 GGUUUCUACCGAUUUAUUGCUACCCAGUUCUACACGCUCCAGCGUUCUUUAUGAG 2835  
Qy 1696 catgtctgaaagatatgacactgaagctttggagaatatagttactatagtgtcataaca 1755  
Db 2836 GGAUUCGCAAAUUCUUGCGAGGCUACUUGCUUACCGACAGUUGCGUUAUAC 2895  
Qy 1756 agggaaaaagcccttcgacaatatgtatgaagatgcttcagactcaacacagaataacc 1815  
Db 2896 CGGAAAGGGAUUGUGGAGAUUUACCAUUGCCAGCUUAGACGCAAGGCGAGCAAGU 2955  
Qy 1816 tcgaagaagcaagatatattggaagctgtatgcatctcaagcagtccttatattt 1875  
Db 2956 CUAAUCUCUUGUUGAGUUCUACUCCUUAUUGCGUGUCGAGGCAUUCUUUUUUC 3015  
Qy 1876 gctatgtgacagaggttctgatttggagcctatctaattcaactgagcgaatgaccca 1935  
Db 3016 UCGUUGCGCUCGUGUUGGAGGAGGAGACUUCUUGGUCACACAGAUUAGAUU 3075  
Qy 1936 gaaggcatgtcatatttcttacttgcaattgacataatgagctatgagcagcaaaacg 1995  
Db 3076 UUGCGUUCUUGUUGUUCUCCGAGAUUUCUUGGUGCUCAUUCGCGGCGCACCGUC 3135  
Qy 1996 ctgcttggctccctgaatatctcaaaagcacaatcgggggtgctgcatctgttgcctg 2055  
Db 3136 UUUUCCUUUGCCAGCAUUGGCGAAGCGAAGCAUUGCGCGCCGCAUUCGACGACUG 3195  
Qy 2056 ttgaaaaaagaacaatatagacagccgcaatcaagaaggaagaaagccagacacatgt 2115  
Db 3196 UUGGACCCGAAAGCCCAAAUUGAUUACUGGUCUAGAGGCGAGAGUUCGAAAGGUG 3255  
Qy 2116 gaaggaaattagagttctcgagaagtccttcttcttctatccaatgtccccaatgtttc 2175  
Db 3256 GAAAGUGAAUUGAAUUAUAGGAGUGGACUUCAGAUUACCGGACCCGACAGACAGCCU 3315  
Qy 2176 atcctcgtgtcttaccctcagttatgaaggaagaagacagtagcatttgtggagc 2235  
Db 3316 GUCCGCGCGGCUUGGACCGUAGCCUGGAAAGCCUGGACAAUUAUUGCGCUUGCGAGCC 3375  
Qy 2236 agcgagcttggaagaagacatctgttcaacttcgcagagacttatgaccgcgtgcaa 2295  
Db 3376 AGCGUUGGUGGAGAGAGUACCAUUCUAGUUGCUUGAGCGCUUUUACGAGUCCGAUUGCC 3435  
Qy 2296 ggaacaagtgcgtttgtatgtgtgtgatacaaaagaattgaattgaatgtgctgcctcc 2355  
Db 3436 GGGUCCAUUCUUGUAGGAGGAGAGACADUAAUUAUAAUUAUUAUUAUUAUUAUUAUUA 3495  
Qy 2356 caaatagcaatcgtctccaagagcgtgtcttcaactgcagatgtgtgagaacatc 2415  
Db 3496 UUUUGUUCACUUGGACGAGGCGGACACUGUACCAAGGCGACCAUACAAGAAAAAUAUC 3555  
Qy 2416 gcctatgtgaacaagccgtgtgtgcatatagatcaaaagaagccgcaaatgca 2475  
Db 3556 UUAUUGUUAUUGUAGAUUAGUACGUAUACCGAGAAUUAUUAUUAUUAUUAUUAUUAUUA 3615  
Qy 2476 gcaaatatcatcttcttatttgaaggtctccctgagaataatacaacaagaattgactg 2535  
Db 3616 GCUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3675

Qy 2536 aaagagcacagcttcttggcgccagaagaacaagaactagctatgtcaaggctctctc 2595  
Db 3676 AAGGAGCAUUDUUGUUGCGCGCCAAAACAGUUGGCGCCGAGCCCUUCU 3735  
Qy 2596 caaaacccaatttatttgttgatgagccactcagccctcgatataatgtacagttgag 2655  
Db 3736 CGGAAUCCCAAUUCUUCUUGCAUUGAGGAGGAGUACGCCUUGACUCCAGUACAGA 3795  
Qy 2656 aagtggttcagatgaccttgataaagccagagcggagaagacatgcttagtgcaat 2715  
Db 3796 AAGGUGUCCAGCGCGCUUUGGAGUCCGCGCCGAGGCGGAAACAAUUCGCGUUGCA 3855  
Qy 2716 cacaagctctgcgaattcagaacgcagattgtatagtggttctgacaaatggaagata 2775  
Db 3856 CACCGACUCCAGCAGAUUCAAAGGCGGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3915  
Qy 2776 aaggacaagaacatcaagaagctcctgagaatacgacataatatttlaagttag 2835  
Db 3916 GUCAAAAGCGAAACGACAGCAUUGGUCAGAAAGGCGGUGACUACGAGCGUGUC 3975  
Qy 2836 aatgcacag 2844  
Db 3976 AACUUGCAG 3984

Search completed: April 22, 2002, 18:39:24  
Job time: 12547 sec

---

...

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:59:36 ; Search time 886.72 Seconds

(Without alignments)  
2761.324 Million cell updates/sec

Title: US-09-873-409-10

Perfect score: 2856  
Sequence: 1 cctattctctcaatctc.....atgcacagtcagtcagta 2856

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: N\_Geneseq\_1101.\*  
2: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1032	36.1	4669	AA052726	Sequence of human
2	1030.2	36.1	3988	AA288973	Human MDR-1 DNA.
3	1030.2	36.1	4349	AAH57442	Human Intestine ce
4	1030.2	36.1	4646	AAQ72872	Human multidrug re
5	1030.2	36.1	4646	AA294738	Human ATP binding
6	1030.2	36.1	4669	AAAT0752	Sequence of human
7	1030.2	36.1	6505	AAAT13394	Hybrid vector PSF-
8	1028.6	36.0	3860	AA249332	Human wild-type mu
9	1028.6	36.0	3860	AA249333	Human G185V mutant
10	1028.6	36.0	4264	AAV65533	Mutated human P-g1
11	1028.6	36.0	4264	AAV65534	Mutated human P-g1

12	1028.6	36.0	8630	AA224042	Retroviral vector
13	1027.2	36.0	3924	AA294742	Human ATP binding
14	1027.2	36.0	3924	AA288973	Human MDR-3 DNA.
15	1027	36.0	4378	AA004522	Multidrug Resistan
16	1027	36.0	8630	AA224041	Retroviral M4 mdr-
17	1026.6	35.9	4279	AAAD0348	Dog P-glycoprotein
18	1026.6	35.9	4279	AAAD0348	Dog P-glycoprotein
19	1026.6	35.9	4279	AAAD0350	Dog P-glycoprotein
20	1025	35.9	4279	AAAD0350	Dog P-glycoprotein
21	1023.4	35.8	4317	AAAD0348	Dog P-glycoprotein
22	1022.2	35.7	4186	AAV32645	Human P-glycoprote
23	1020.6	35.7	4186	AAV32645	Cynomolgous monke
24	1020.6	35.7	4195	AAV32647	Rat multidrug resi
25	967.8	33.9	4369	AA252047	Rat multidrug resi
26	967.8	33.9	4425	AA252048	Rat multidrug resi
27	953.2	33.4	4233	AA250198	Rat mdr1b2 (multis
28	950.2	33.3	4233	AA250198	Rat mdr1b2 multidr
29	950.2	33.3	4233	AA250198	Murine multidrug r
30	949.6	33.2	4189	AA249334	Murine multidrug r
31	949.6	33.2	4313	AA038950	Mouse multidrug re
32	856	30.0	2726	AAQ70907	Multidrug-resistan
33	856	30.0	2726	AAQ70916	Multidrug-resistan
34	856	30.0	2726	AAQ70916	Multidrug-resistan
35	669.8	23.5	4776	AA249744	Human ATP binding
36	478	16.7	4175	AAV69392	H. contortus PGP-A
37	425	14.9	4051	AA090019	A. thaliana ATPAC
38	414.6	14.5	4002	AA85823	CDNA encoding mult
39	409.8	14.3	4102	AA45942	Arabidopsis thalia
40	383.8	13.4	2698	AAV69394	H. contortus PGP-A
41	374	13.1	4047	AA60201	CDNA encoding the
42	329.6	11.5	3512	AAV69395	H. contortus PGP-O
43	315.6	11.1	8777	AAV69399	P. chrysogenum ABC
44	315.6	11.1	8777	AAV69399	P. chrysogenum ABC
45	315	11.0	3133	AAV69398	P. chrysogenum ABC

#### ALIGNMENTS

RESULT	ID	AA052726 standard; CDNA: 4669 BP.	Location/Qualifiers
1	AA052726		
AC	AA052726		
DT	24-JUN-1994	(first entry)	
DE		Sequence of human multi-drug resistant gene mdr1.	
DE		Sequence of human multi-drug resistant gene; mdr1; ss.	
KW		P-glycoprotein; multi-drug resistant gene; mdr1; ss.	
OS		Homo sapiens.	
FT	Key	Location/Qualifiers	
FT	CDS	424..4267	
FT		/*tag= a	
XX	W09324613-A.		
XX	09-DEC-1993.		
XX	14-MAY-1993;	93WO-US04707.	
XX	22-MAY-1992;	92US-0887712.	
XX	(GENE-) GENETIC THERAPY INC.		
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
XX	McDonagh KT, Nienhuis A, Tolstoshev P;		
XX	WPI; 1993-405805/50.		
XX	P-PSDB; AAR44297.		



```
OY 2044 ctgttgcctgtgttgaaagaacccaatatagacagccgagtcacgaagaagaaag 2103
DB 3446 atcaccatgatcatctgaaacacccctttagtaccagctacagacaggaagcctaaag 3505
OY 2104 ccagacacatgtgaaaggaaatttagatcttcgaagatctcttcttcacatcagtcgc 2163
DB 3506 ccgaacacatcttgaaaggaaatttagatcttcgaagatctcttcttcacatcaccgca 3565
OY 2164 ccagatgttctacatccctcgtgtgtatccctcagatattgagcgaagaaagacagtaaga 2223
DB 3566 ccgacatcccatgtcttcaagagactgagccctggaagtgaagaaggaagccagcgcgtgc 3625
OY 2224 ttgtgtgggagagcagcgtgtgtgggaaagcactctgttcaactctctgagaactttat 2283
DB 3626 cgtgtgaggaagcagtcgtgtgtgggaaagcagatgtgtccagctctctcagtcggttctac 3685
OY 2284 gacccctgtgcaagcagacagtcgtgtgttgatgtgtgacatcaaaagaaatgtgattacag 2343
DB 3686 gacccctgtgcaagcagacagtcgtgtgttgatgtgtgacatcaaaagaaatgtgattacag 3745
OY 2344 tggctccgttcccaaatagcaatcgcttctcacaagagcctgtgtcttcaactgcagcatt 2403
DB 3746 tggctccgtgcaacacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3805
OY 2404 gctggaacacatgcctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3865
DB 3806 gctggaacacatgcctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3865
OY 2464 gctggaacacatgcctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3925
DB 3866 gctggaacacatgcctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3925
OY 2524 caagtgtgactgaaagagcagacgttctgcgagccagaaacaaagactgatttga 2583
DB 3926 aagatgagagacaaagaaactcagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3985
OY 2584 aaggtctcttcccaaaaacccaatttatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2643
DB 3986 cgtgtccctgtgtgacagcctcatttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4045
OY 2644 aatgcacgtgaaagagtggttgcagatccctctgataaaagcagagcaggaagagacatc 2703
DB 4046 acagaaagtgaaagagtggttgcagatccctctgataaaagcagagcaggaagagcagac 4105
OY 2704 ctatgtgtcactccacagcgtctctcgaattcgaacgcagatttgaattgtgttcac 2763
DB 4106 atgtgtatgtctcaccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4165
OY 2764 aatggaagaataaagaaagaaactcacaagaagcctctcgaagaatcgaagacatattat 2823
DB 4166 aatggaagaataaagaaagaaactcacaagaagcctcgaagaatcgaagacatattat 4225
OY 2824 tttaagttagtgaatgcacagtcag 2848
DB 4226 ttctcaatgtgtcagtgctcaggtcg 4250
```

```
RESULT 2
AAZ88973
ID AAZ88973 standard; DNA; 3988 BP.
```

```
XX AAZ88973;
```

```
DT 30-MAY-2000 (first entry)
```

```
DE Human MDR-1 DNA.
```

```
XX Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;
```

```
KW stomach; colon; ulcerative colitis; pouchitis; large intestine; human;
```

```
KW inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;
```

```
KW ileum; anti-ulcer; anti-inflammatory; cytosolic; MDR-1; ss.
```

```
OS Homo sapiens.
XX WO200007577-A2.
PN 17-FEB-2000.
XX 06-AUG-1999; 99WO-DE02426.
XX 06-AUG-1998; 98DE-1035526.
XX 15-DEC-1998; 98DE-1057750.
XX (STRE/) STREMMEL W.
XX Streammel W.
XX WPI: 2000-195439/17.
DR Using phosphatidylcholine having mucous membrane protective activity to
XX prevent or treat large intestinal diseases, such as ulcerative colitis
XX or pouchitis.
XX Disclosure: Page 17-22; 22pp; German.
XX
XX This invention describes a novel medicament containing a therapeutically
XX effective amount of phosphatidylcholine to treat diseases, where
XX phosphatidylcholine has an advantageous mucous membrane protective
XX activity in the large intestine. The invention describes (1) a method
XX to regulate presence or absence of MDR(multidrug resistance)-3
XX transcription in a sample from patients with suspicion of stomach and
XX colon diseases, where treatment with phosphatidylcholine is indicated;
XX (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large
XX intestinal inflammation, Crohn's disease, diverse colitis, infectious
XX enteritis/colitis, inflammation through X-ray treatment, antibiotics,
XX chemotherapeutics, drugs or chemicals or large intestinal carcinoma,
XX contains MDR3-specific primers to detect mutations in the MDR3 gene
XX family and their interactions partners (in particular transcription
XX factor); and (3) a method to manufacture a medicament for local gene
XX therapy in the ileum with the ideal MDR3-analogous genes for patients
XX with missing or decreased expression of MDR3-analogous proteins. The
XX products of the invention have anti-ulcer, anti-inflammatory and
XX cytosolic activity. The orally administered delayed-release form of
XX phosphatidylcholine prevents premature resorption and provides for
XX targeted release in the lower section of the small or large intestine.
XX This sequence encodes the human MDR-1 protein described in the method of
XX the invention.
XX
XX Sequence 3988 BP; 1173 A; 763 C; 988 G; 1064 T; 0 other;
SQ
Query Match 36.1%; Score 1030.2; DB 21; Length 3988;
Best Local Similarity 62.2%; Pred. No. 2.4e-263;
Matches 1695; Conservative 0; Mismatches 986; Indels 42; Gaps 3;
OY 166 attattgataaagaaacccagatatagataactttccacagcctgcatataaactgaatcc 225
DB 1168 ataattgataaataagcaagaatttgacagctatttgaagaagtgagcacaacacagataat 1227
OY 226 atagaaggaactgtggaatttaaaaatgttcttccaattatcacaagaccatctatc 285
DB 1228 attaagggaatttgaaattcagaatgttcaacttcaagttacccatctcgaagaagaagt 1287
OY 286 aagattctgaaaggtctgcatatcgaatgaatgctggaagaagcgtgcctgtgtgc 345
DB 1288 aagattctgaaaggtctgcatatcgaatgaatgctggaagaagcgtgcctgtgtgc 1347
OY 346 ctcaatgagcagtggaagagctagcgtgacagctctcgaagaatttatatccagat 405
DB 1348 aacagtggtgtgtggaagagcagcagcagcgtgacagcgtgacagcgtgacagcagc 1407
OY 406 gatgcttatacatgtgtgataagaatcacaagcagcagcgtttaaattgtcgcgcatatcga 465
DB 1408 gagggagatgtcagtggttgaatgacagagatatagacacataatgtaagtgttctcag 1467
```

QY	466	gaccatattggaagtggttaagtcagaagcctgttttgcctcgggacacacatcaagtacaacat	525
Db	1468	gaacatctgtgtgtgtgagtcagaagacacctgattgttgcaccacagataagctlaaagac	1527
QY	526	atcaagtatgagcagagatgatagtatgaactatgaagatgagagagcagcaagggaagca	585
Db	1528	atctcatctgagccgggaanaatttcacacatgatagtatgaaagcgttccaaggagacc	1587
QY	586	aatgtatgatatttcaacatgtagtctccataataatttaacattacatttgaaggaaaaa	645
Db	1588	aattgcatatgacttatactatcatgaaactgtcctataaatttgaaacacctgtgtgagagaga	1647
QY	646	gaagcttcaaatgagtggagggcagaacaagagatcgcgaattgtctgtcgttgaattcgaa	705
Db	1648	ggggccagcttgatggtgtggtggcagaagcagagatcgcattgcacgttccctgttcgc	1707
QY	706	aaccccaagatctgattttaatgtaggttaagtcgtccctgtgattccagaagcagaatga	765
Db	1708	aaccaccaagatcctcctgtctgtatgtaggcacagctcagctctgtgacaagaagcgaagaa	1767
QY	766	gctgttccaagctgcgaactgtgagaagcggagcaaggtccgacatacaactcgtgtgaagcac	825
Db	1768	gtgtgtccaagtggtgctctgtgataagccgaaaaaggtcggaaaccaattgtgataagctcat	1827
QY	826	cgactttccactatttcgaagtgtcagatttgatgttgaccctaaagaaatgaaatgtctgctg	885
Db	1828	cgtttcttcaagcttgctgaatgctgaagtcacatcgtctgttctgatgtgagatcattgtg	1887
QY	886	gagaaaggagcacatgtctgaactaatgtgcacaaacgggtctatattatccactgtgagt	945
Db	1888	gagaaaggaaatcatatgtatgaacttcaatgaaagagaaaggaattacttcaaacattgtcacaa	1947
QY	946	tcaacagatatataaanaagctgtatgacacagatgagtcagtcacatgacatactta-----	997
Db	1948	atgcgaacagcaggaataatgatagttgaattcagaanaatgcgcgtgtatgataicccaaagtga	2007
QY	998	-----ctgaagaagaagaccacactacttccctctgcacctctgtgaagaagatca---	1045
Db	2008	attgatgaccttggaanaatgtcttcaaatgatattcaagatccacgtcttaataagaanaagatca	2067
QY	1046	-----agtaagacttattgacaagcgtgaagaaatcaccacacatct	1086
Db	2068	actcgttaggaatgtccgttgatctacaaagcccaagacagaagaagcttgaaccaaagaagct	2127
QY	1087	aagagagaatgctctctctgtagcctctcctaattaaaaattttaaagtaaaacaagcgtgaa	1146
Db	2128	ctggaatgaaagatatacctccaagttcctcttggagaatgataatgaagcctaaatttaactgaa	2187
QY	1147	tggcctcttctgtgtctcgtggagcaatgtgcttctgtctaaatggaactgttcatcagaatga	1206
Db	2188	tggcctatttctgtgtgtgtatttctgtgcacattaaatggaagcctgcgtcaacacagca	2247
QY	1207	ttttccatcattcttggaaaaaattatracatcgttttgaaataatgataaac---aca	1263
Db	2248	tttgcataatatttccaagaatataaggggtttttaaagaagaattgtcagatcctgaaaca	2307
QY	1264	ttaaagatgatgagagaattatattccatgataatctgcatttttgggttgatattgtctt	1323
Db	2308	aaagcagaagaatgataactgttttcaactatgtttctaagccttgaattatttcttt	2367
QY	1324	gtcagattattcaltcagcaggaattatttcaaggcagacaggggaatttaaagatgaga	1383
Db	2368	attacattttccttccagagtttcaacttttgcagaagctgtgaagatcctctacacaaacgg	2427
QY	1384	ttaagacactgtgccttccaagaagccatgttatatacagatatctgcctgtttgataagaag	1443
Db	2428	cttcgataacatgtttcttcgaatccatgtcgcagaagaagatgtagtggattgtgagaccct	2487
QY	1444	gaaaaacgcaaggaagccttgaacaacaatatctagccataagatatagacaacaatccaaga	1503
Db	2488	aaaaaacacacatgtgacatgttgcataccagcgtcgcacaatgatagtctgtctcaagttaaagg	2547
QY	1504	gcaacaggttccaggaattgtgggtcttlaacacaaatgnaactaacaatgtaggaacttcagtt	1563

Db	2548	gctatagttccaggctctgctctgtaattacccgaaatatacgaaatctctggacagata	2607
Qy	1564	atcatcttccttatatagatgagtggsagatgacatcttcgatctcgaagtattgctccagta	1623
Db	2608	atatactctcatcattatggttggcaacaacaacgttactctttagacatattgtaccatc	2667
Qy	1624	cttgcgtgacagaatagatctgaacccgagcaatgacgttggatttggccaacaagaatga	1683
Db	2668	attgcataatagacagatgttcttgaataatgtaatgtctgcggcaacgctctgaagaata	2727
Qy	1684	caagaacttaagcagctgtctggaagaatagacaacttgaaagctcttggagaatatacgtata	1743
Db	2728	aaagaactagaaggtgctctggaaagacgcttactcgaacatagaaaacttcggaacct	2787
Qy	1744	gtgtcatatacaaaagggaanaaaagccttcgagcaaatgtatgaagagatgcttcagatcaa	1803
Db	2788	gttctcttgactcagagcgagaagtgttgacaatatgtatgctcgaagatttgcagatcaa	2847
Qy	1804	cacagaatactctgaaagaagcagcagattattggagaagctgttatgactatgaacagtc	1863
Db	2848	tacagaataactcttggaaagaagcaacaactcttggaaatatacttccccaacccagca	2907
Qy	1864	tttatataatttgcctatgcagcagaggttccgaatttggagcctataatcaagcttga	1923
Db	2908	atgagatgtatttccctatgctcgatgttcccggttggagcctacttggtcacataaa	2967
Qy	1924	cgaaagacccagagggagatgttcataagttttaaactgcaattgcaatatgagctatggcc	1983
Db	2968	cccatgagcttggagatgctgtctgtatgatttccagcgtgtcttcttggtcacatggcc	3027
Qy	1984	atcggaaaaagcctcgttcttggcctccgaatctcacaagaacccaatccggggcgctgc	2043
Db	3028	gtggggagcaagtcatcattgtgctccctgactatgcaaaagccaataatcagcagccac	3087
Qy	2044	ctggttgccttggtaagaagaacccaatcatagacagccgagctcaagaagaaggaaaag	2103
Db	3088	atcatcatgcatcattgaaaacaccccttggattgacagcctcaagccaggaagcctaag	3147
Qy	2104	ccagaacatgctgaaaggaattagatcttcgagaagctcctcttctctatccatgtgc	2163
Db	3148	ccgaacacatggaagaaatgcatcatttggtagaagttgtatctcaactatccaccgga	3207
Qy	2164	ccaagatgtttatcactctccgctgagcttatccctcgaatatatgagggagaagaacagta	2223
Db	3208	ccggaacatcccaatgctctcaaggagcctgagccttgaagtgaaagaaggccaagcctgct	3267
Qy	2224	tttggtagagagagagcgtctgctggaagaagcactctctgttcaacttctgacagacttat	2283
Db	3268	ctgggttgagagagcagtgctgtgtgggaagaagcaagtgctccagctctggaagcgtttac	3327
Qy	2284	gaaccctgcaaggaacagtgctgttctgaatgagtgatgagtgatgaag	2343
Db	3328	gaacccttgcagaggaagtctgctgttgaatgcaagaagaataagcactgaaatgtccg	3387
Qy	2344	tggcttcgttcccaaatgagcaatctgttccctcaagaagcctgtgcttccaactgcagcat	2403
Db	3388	tggcttcgagcacaccctgggggcatctgtctgcaccagaagccatcctgtttgactgcagcat	3447
Qy	2404	gtcgaagaacatgcctatgtatgacaacagccgtgtggtgacattagatgaaacaaaga	2465
Db	3448	gtcgaagaacatgtcctatgtagagacaacagccgggtgtgtcacaaggaagatgttgaag	3507
Qy	2464	gtccgcaaatgcaagcaatatccattctttatgtgaaggtctccctgagaataataacaca	2522
Db	3508	gcaagcaaaaggagccaacatacatgctcttcacgtacgtacgtctaaataataagcact	3567
Qy	2524	caagtttgacctgaagagagcaagcttctctggcgccaagaacaagaactagatttga	2584
Db	3568	aaagttagagagacaagaactatagctctctctgtggtgcaggaacaacgcatctccatgct	3622
Qy	2584	agggtcttctccaaaacccaattttatgtgtgagtgagggcaacttaagccctgact	2642

Db 3628 cgtgcccttgtagacagcccatatttgccttttgatgaagccagctcagctcgat 3687  
Oy 2644 atgacagtggaagagtggttcacagatgccttataaagccagagacggaagcagtc 2703  
Db 3668 acagaaatgaaaggtgtccaaagaagccttgacaaagcaggaagagccgcacccgc 3747  
Oy 2704 ctatgtgcatcctcagagctctgcgaatcagacagcagattgatgtgttcgac 2763  
Db 3748 atgtgtatgtctcaccgctgtccaccatcagaatgcagctcatatgtgtttcag 3807  
Oy 2764 aatggaagaataaaggaacgaagcatcacaagaagctcctggaagaatcgacaatat 2823  
Db 3808 aatgcagacagtcagcagcatgcacgcacagctgcgtgcgacagaagcagcatcat 3867  
Oy 2824 tttaagttagtgaatgcacagtcag 2848  
Db 3868 tttaacatgctcagtgctcagagctg 3892

## RESULT 3

AAH57442

ID AAH57442 standard; cDNA; 4349 BP.

XX

AAH57442;

XX

10-SEP-2001 (first entry)

DT

Human intestine cell specific cDNA sequence S80 ID NO:282.

DE

XX

Human; tissue specific; diagnosis; brain; heart; skeletal muscle;

KW

lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;

KW

metabolic disease; developmental disease; cytostatic; immunomodulatory;

KW

neuroprotective; gene therapy; cancer; Immunopathology; neuropathology.

KW

Homo sapiens.

OS

MO200132927-A2.

PN

10-MAY-2001.

PD

02-NOV-2000; 2000WO-US30396.

PF

04-NOV-1999; 99US-0163508.

PR

(INCY-) INCYTE GENOMICS INC.

PA

Sornasse T, Sellhammer JJ, Watson GA;

PI

WPI: 2001-291057/30.

DR

New cell and tissue specific polynucleotides useful for diagnosis,  
prognosis or monitoring of treatments for disorders where the gene is  
associated with a cancer, immunopathology or neuropathology -  
Claim 1; Page 207-208; 327pp; English.

PS

XX

AAH57461 to AAH57576 represent cell and tissue specific polynucleotide  
sequences (I). (I) can have cytosolic, immunomodulatory and  
neuroprotective activities, and can be used in gene therapy. (I) and  
proteins (II) encoded by them are used in high throughput screening  
assays to select DNA molecules, RNA molecules, peptide nucleic acids,  
mimetics, peptides, proteins, agonists, antagonists, antibodies or  
their fragments, immunoglobulins, inhibitors, drug compounds and  
pharmaceutical agents. Expression of (I) in a sample indicates the  
differentiation of embryonic stem cells into a tissue selected from  
brain, heart, kidney, liver, lung, skeletal muscle or pancreatic  
tissues. (I) and (II) are used to produce an expression profile that  
defines a metabolic or developmental process, treatment, condition,  
disease or disorder. The gene profile can be used for diagnosis,  
prognosis or monitoring of treatments and for investigating a  
predisposition to a disorder where the gene is associated with a  
cancer, immunopathology or neuropathology.

CC

XX

SQ Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Query Match 36.18; Score 1030.2; DB 22; Length 4349;

Best Local Similarity 62.28; Pred. No. 2,56-263;

Matches 1695; Conservative 0; Mismatches 988; Indels 42; Gaps 3;

Oy 166 attattgtaagaaccagatagataactttccacagcttgatataaactgaatcc 225  
Db 1227 ataattgataaaccagatattgacagctcattcgaagatgggacaacaacataat 1286  
Oy 226 atagaagaactgtggaatttaaaatgttcttcaattatccatcaaacatctc 285  
Db 1287 attaagsgaaatttggaattcaagaatgttcaactcagttaccatcccgaaagaagt 1346  
Oy 286 aagattctgaagagctgaatctcagaattgaagtcgagagacagctgcgtgctg 345  
Db 1347 aagattctgaagagagctgaatctcagaattgaagtcgagagacagctgcgtgctg 1406  
Oy 346 ctcaatgacagtggaagaagtagcgttaagtcagcttcctgcagaggttatatccggt 405  
Db 1407 aacagtgctgtyggaagaagacacacagctgtagcagaggtctctatgacccaca 1466  
Oy 406 gatgcttatcatgtgtgtagatgagatcacatcagagctttaaattgtgcgcatatcga 465  
Db 1467 gagggagatgctcagtgatgtgtagagacagatatatgacacataatgtaagtttctag 1526  
Oy 466 gaccatattgagtggttagtcaagaagcctgtttgttcggagaccacatgaatacaat 525  
Db 1527 gaattcatgtgtgtgtgtagtcaagaacctgtattgttcgacacagatgctgaagaa 1586  
Oy 526 atcaagatgagcagagatgtagtactgtgaagaagatggagagagccagcagaagaa 585  
Db 1587 atcgtctatggccgtgaaagagtcacatgtagtgaattgaaagcgtccaaagagcc 1646  
Oy 586 aatgctatgtattatcatcagatgagattcctaataattatatactgtgagggaa 645  
Db 1647 aatgctatgtattatcatcagatgagattcctaataattatatactgtgagggaa 1706  
Oy 646 ggaagctcaaatgagtgagagcagaagaacagagatcgcaattgtcgtgctttagtga 705  
Db 1707 gggagccagatgagtgagtgagcagagagagatcgcaattgtcgtgctttagtga 1766  
Oy 706 aacccaagattctgattttagatgagctcgttcgctcggttcaagaagcaga 765  
Db 1767 aacccaagattctcctcgtcgtgagatgagcagcagctcgttgaacacagaagcaga 1826  
Oy 766 gctgtcagctcagctggaagaagcagcagaagtcgagactacatcgtgtgtagcacac 825  
Db 1827 gtgtgtcagtggtcctgtgataaagccagaagaagtcgacacacattgtgataagctat 1886  
Oy 826 cgaacttactactcgaagtgcagattgtatgtgacccataaagatgtaagtgtgcg 885  
Db 1887 cgttgcctacagttcgtgaatgtagcgtcatcgtgtgttcgagatgtaggagtcattg 1946  
Oy 886 gagaagaagcacaatgcctgcaactaatgycacaagaggtcatattatcactgtgtg 945  
Db 1947 gagaagaagcacaatgcctgcaactaatgycacaagaggtcatattatcactgtgtg 1006  
Oy 946 tcaacaggtatlaaaagcctgataaagaagtgagtcataatgacatatca----- 997  
Db 2007 atgcagacagcagaagaatggaattgaataaataagcagctgtagatcccaagaagtga 2066  
Oy 998 -----ctgaagaagaagcacaactcctcctcgtcagctgtggaagcagatca--- 1045  
Db 2067 atgtgtcctggaagtgtctcacaatgataagatccagttcactaataagaagaagatca 2126  
Oy 1046 -----agtcagactcattgtcagaagcagctgtagatcccaactc 1086  
Db 2127 actgttaggaatgtcgtgtgatacacaagccagaagcagaagcagctagtagcagaagct 2186  
Oy 1087 aaagagataaagcttccctggaagctcctctatataaattttaaagttaaagaagcgtgaa 1146



Db 2187 ctgagatgaagatgaacctcagcttcttccttttgaggatattgaagctaaatttaactgaa 2246  
QY 1147 tggccttttggtctctgaggacattggtcttgcgttcaatgaactgttaccagta 1206  
Db 2247 tggcctattttgttggatgtatttttgcacattataatgaggccgcagca 2306  
QY 1207 ttttccatctcttggaataaaltataacacgttttggaataatgaataaac--aca 1263  
Db 2307 ttgcataataatttcaagaattatagggttttttcaagaatttgatgacctgaaca 2366  
QY 1264 ttaagatgagtgcagaattatttccatgattttgctattttgggtgtattttgctt 1323  
Db 2367 aaacgaaatagatgaactgttcttcaattgttctagcccttggaattatttctt 2426  
QY 1324 gtacgtatttctatgcagggattatttcaagcagagcagggaaatttcaacgataga 1383  
Db 2427 attacattttcttccctcagggttccacatttggcaaaacttggagagatccccaagcgg 2486  
QY 1384 ttaagacatttgccttcaaaagccattatlatcagaatttgccttggtttgataaag 1443  
Db 2487 ctccgatacatgttttccgatacctatgctcagaagaagtgtgattgttgaagccct 2546  
QY 1444 gaaacacgcagagagcttgcagacaataattgccaatagatatagacaataattcaagga 1503  
Db 2547 aaaaacacccctggaagatgactgaacacagctccgcaatgagtctgcagatgaaggg 2606  
QY 1504 gcaacaggttccagagatttgccttcttacaacaataatcaactaacaaggttccagtt 1563  
Db 2607 gctataggttccagagcttgcgttaattaccagaataatagcaaatcttggacagata 2666  
QY 1564 atcatttcttataatgtatgtggagatgacatctcgtattctgagattgttccagta 1623  
Db 2667 attataccttcaactatgttggcaactaacaactgttactcttagaactgttccatc 2726  
QY 1624 ctggcgtgacaggaatgatttgaacacgcagcaatgactgatttgcacaagaataag 1683  
Db 2727 attgcagtaagagaggttgttgaataatgttgccttgcagacaagcactgcgaataag 2786  
QY 1684 caaagaacttaagcactgtctggaataaagatagcaactgaacttggagaatacagtaata 1743  
Db 2787 aaagaactagaaggttgcgtggaagatcgtctacatgaagcaataagaaacttccgaagct 2846  
QY 1744 gtgtcattacaagaaggaataagccttcgagcaaatgtatgaagaagatgctcagaactaa 1803  
Db 2847 gtcttcttgactcagagcaggaagtttgaacataatgatactgacagatttgcagtaaca 2906  
QY 1804 cacaagaatacctcgaagaagaagcagaattatttgaagagctttagatttaccagctgac 1863  
Db 2907 tacaagaactccttgaaggaagacacacacttggaaattacatttccatccacagga 2966  
QY 1864 ttataataatttgccatagcagcagggttccgacttggagccataatttaactcaagctgga 1923  
Db 2967 atgattgatttttctcctatgctgagatgttccggttggagcctacttgcgtgcacataa 3026  
QY 1924 cgaatgaccccaagagggcagtctcatagtttttactgcaattgcataatgagctatggcc 1983  
Db 3027 ctctatgagcttggagatgttctgttgaatttccagctgttgccttgggtgcacatggcc 3086  
QY 1984 atcggaanaagcctgttggctccggaataatccaaagcaaatcgggggctggcagat 2043  
Db 3087 gtggggcaagtcagttcaattgtccttgacatgacaaagccaaataatcagcagccac 3146  
QY 2044 ctgttgccttcttgaagaagaacaataatagaacagcccgatcgaaggaaggaagaag 2103  
Db 3147 atcatatgattcatgaaaaaaccccttggattgacagctacacaggaagcgtctaatg 3206  
QY 2104 ccagacacatgtgaaggaatttagagtttgcagaagctcttcttccatcatgtcgc 2163  
Db 3207 ccgaacacatgtgaaggaatgacatlttggatgtatcattcaactacccaaccca 3266  
QY 2164 ccaagatgttccatccctccggtgcttaccctcagatatttggcaggaagaacagtagca 2223  
Db 3267 ccgagacatcccatgtctcaggaagctgagccttgaggtgaaaggaagggccagacgctgct 3326

QY 2224 ttgttggggagcagcggtctgttgggaagaagcacttctgttcaacttctgcagactttat 2283  
Db 3327 ctgttggcagcagctgtctgttggagaagacagatgttccagctctcgttgcgtttac 3386  
QY 2284 gaccctgtcagaagcaagcagctgttggatgtgtgtgagtcgaagaatgagtagcaag 2343  
Db 3387 gacccttgcaggaagagctgtctgttgaatgcagaagaataaagcagctgaatttccag 3446  
QY 2344 tggctccgttcccaatagatcgttctcctcaagagccttgccttcaactgcagact 2403  
Db 3447 tggctccgagacacacccgtggcatcgttcccaaggagccacttctgttgcagcatt 3506  
QY 2404 gctgagaacatgcctatgttgaacaacagcgttgccttccatagatgatacaagaa 2463  
Db 3507 gctgagacacttgcctatgtgaacaacagcgggtgtgttcaagaagaagatcgtgaag 3566  
QY 2464 gccgcaatgcagcaaatatccatttcttatttgaaggttcccttgaagaatacaaca 2523  
Db 3567 gcagcaaggaaggccaacatcacatgccttccatcgaagtcagctgcttaataataagact 3626  
QY 2524 caagtgcactgaagaagcagcagcttctcgcgcgcaggaacaagaactagctatgca 2583  
Db 3627 aaagtgaagaaagaagaactcagctctcgttgcagagaacaagcagctgcacatgct 3686  
QY 2584 agggcttctccaaaaccacaatttattgttgaatgagcacttaagccctgact 2643  
Db 3687 cgtgcccttgttagacagcctcatatttgccttattgaatgaacagcagctgactgact 3746  
QY 2644 aatgacagtagaagctgttgcagcagcttcttgaatgaagaagcaggaaggaagactg 2703  
Db 3747 acagaagaaggaaggttgcctcagaagaagccttgcagacaagcagagaagcgcagactgc 3806  
QY 2704 ctagtgtcactcacaagcctctgtcaattcagaacgcagatgtgattgttctgcac 2763  
Db 3807 atgtgttgcctcaccgcctgcacacatccagatgcagacttaatagctgttccag 3866  
QY 2764 aatggaagataaagaagaagcaacccatccatcagaagctccttggaaatcgcacatat 2823  
Db 3867 aatgacagatcagaagcagctgtgcacgcatcagaagctgtgtgcagagaaggaactat 3926  
QY 2824 tttaagttagtaagtcacagctcag 2848  
Db 3927 ttccaatgctcaggtccagcgtg 3951

RESULT 4  
AAQ72872  
ID AAQ72872 standard; cDNA; 4646 BP.  
XX  
AC AAQ72872;  
XX  
DT 21-JUN-1995 (first entry)  
XX  
DE Human multidrug resistance gene (MRD-1) encoding a P-glycoprotein.  
XX  
KW Human; tyrosinase; compound; hair; follicle; liposome; lipophilic;  
KW lipophilic; stratum corneum; melanin; hair dye; glycoprotein; alopecia;  
KW chemotherapy; transforming growth factor; growth stimulant; aromatase;  
KW cyclosporin A; macromolecule; polymer; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FH CDS 425..4267  
FT /\*tag= a  
FT /product= P-glycoprotein  
XX  
XX MO9422468-A.  
XX  
XX PD 13-OCT-1994.  
XX  
XX 01-APR-1994; 94KW-US03634.

XX 02-APR-1993: 93US-0041553.  
PR 13-JAN-1994: 94US-0181471.  
XX (ANTI-) ANTICANCER INC.  
XX LI L, Lishko VK.  
XX WPI: 1994-332816/41.  
DR P-PSDB: AA072872.  
XX  
PT Liposomes for delivering protein, nucleic acid etc. to hair  
PT follicles - e.g. to restore hair colour, prevent hair loss during  
PT chemotherapy, stimulate hair growth etc.  
XX  
PS Claim 8: Page 70-77; 100pp; English.  
XX  
CC The nucleotide sequence of the human multidrug resistance (MDR-1) gene  
CC which encodes a P-glycoprotein. This is an example of a compound which  
CC can be delivered to hair follicles via a novel liposome composition.  
CC The liposomes are comprised of a lipophilic or lipophobic compound which  
CC will selectively target the hair follicle (via the stratum corneum)  
CC without damaging or unwanted effects on cells outside the follicle.  
CC Compounds e.g. P-glycoprotein, can be delivered to the hair follicle for  
CC treatment of chemotherapy-induced alopecia. Other compounds targeted at  
CC hair follicles can include: tyrosinase (or the DNA encoding it -  
CC AA072871), melanin or hair dyes (to restore hair colour or condition);  
CC human transforming growth factor-alpha (AA072873) (for reversal of wavy  
CC hair); hair growth stimulants (e.g. cyclosporin A or aramatase) or  
CC antisense sequences. The method allows compounds (e.g. macromolecules  
CC or polymers), which would not normally reach the hair follicles, to be  
CC delivered to these target areas.  
XX  
SO Sequence 4646 BP; 1371 A; 893 C; 1128 G; 1254 T; 0 other:

Query Match 36.1%; Score 1030.2; DB 15; Length 4646;  
Best Local Similarity 62.2%; Pred. No. 2.6e-263;  
Matches 1695; Conservative 0; Mismatches 988; Indels 42; Gaps 3;

QY 166 attatctgaagaacacagatataataacttccacagctggaataaacttgatcc 225  
DB 1526 ataattgtaataagccagatcttgacagctatcgaagagtgggcacaaccagataat 1585  
QY 226 atagaagaactgtggaattaaatgttcttcaattatccatcaagaaccatctatc 285  
DB 1586 attaagggaatttgacatgaatgttccattcagttaccatctcgaataagaagt 1645  
QY 286 aagatctggaagctctaactcagaatttaagcttgagagacagctggctgtgctg 345  
DB 1646 aagatcttgaagggcctgaacctggaagctgcagagtgggcagcgtggtgttga 1705  
QY 346 ctcaatgtcagtggaagaagtagcgttagctccagcttcgcaaggtttatcatcagat 405  
DB 1706 aacagtgctgtgggaagagcacaacagctccagctgcatgcaaggtctctatgacccaca 1765  
QY 406 gatgcttcatcagtggtgagatgagaatgacatcagagctttaaattgctggcgcattatcga 465  
DB 1766 gagggatggtctcagtggttgcagacagatatagaccataaatgtaagtcttctacg 1825  
QY 466 gaccatattggtgtgttagtcaagaagcctgttgttcggagacacatcagtaaat 525  
DB 1826 gaaatcatgtgtgtgtgagtgcaaggaacctgtatgttgcgcaacacagatagctgaaac 1885  
QY 536 atcaaatgtagcagagatgagtgatgctgataagaagatgagcagcgaagggaagca 585  
DB 1886 atctgctatgscgtggaanaatgtccacatgtagtgatgtagaagaagcgttcaaggaagcc 1945  
QY 586 aatggtatgattatcatatgagatgttccctaataattatcatcattgtaggggaaaaa 645  
DB 1946 aatgctatgattatcatatgagatgttccctaataattgacacccgttltgagagaga 2005  
QY 646 ggaagtcaaatgagtggaggggcagaagaacagagatcgcaattcgccgtgtcttctga 705

DB 2006 gggggccagattgagtgtgtggcagagaagagatcgcgcatcttgcaagcgtgttctgc 2065  
QY 706 aacccaagatctgattttagatgagctacgtctccctggaattcagaagaagatca 765  
DB 2066 aacccaagatctctctctgtgtaggagcagctgaaccttggaacagaagaagca 2125  
QY 766 gctgttcagtcgacgtgaggaagcagcgaaggtcggaactacatctgtgtgtagaac 825  
DB 2126 gtgttcaggtgtgcttgagtaagagccagaanaagctgcgacacatctgtatgacct 2185  
QY 826 cgaacttctactatcgaagtgcaagatttgatgtgacacctaaagaagatgctgtgc 885  
DB 2186 cgttgtctacagtcgtcaatgctgacgtcatcgctgttctcgatgagggagctgtgt 2245  
QY 886 gagaagaagcagatgctgaactaattgcaaaaagagctatattatctactgtgag 945  
DB 2246 gagaagaagaaatcagatgaaactcagaaagaagacatcttacttcaaaactgttca 2305  
QY 946 tcacagatataaaaaagcgtgatgacaagatgaggtcaatgacatattca----- 997  
DB 2306 atgcagacagcaagaatgaaagttgaaatgaaatgcaagctgatgaaatccaaagtga 2365  
QY 998 -----ctgaagaagaagcaactcaactcctctgcaactgttgaagaacatca--- 1045  
DB 2366 attgattgcttggaagtgtcttcaaatgattcaaatcagctcatatgaagaagaatca 2425  
QY 1046 -----agtcagaactcattgcaagaagctgaggaatccaccatct 1086  
DB 2426 actcgtagagtgctcgtgagacacaagcccaagcagaagaagcttagaacaagaagct 2485  
QY 1087 aagaagataagcttcctcgaagctcctctatataaaatttaagtttaacaagcttga 1146  
DB 2486 ctgagtaagaagataccctcccaagcttcccttcttggaagatatagaagctaaattca 2545  
QY 1147 tggccttgtgtctgtcggagacattggtcttctgtcctaattgaagacgtgtcatcaca 1206  
DB 2546 tggccttatttgtgtgtgtgtatttcttgccatataaagggcgtcgaacaagca 2605  
QY 1207 ttccatcatcttgcgaaaaattataacacatggttggaaataatgataaaac---aca 1263  
DB 2606 ttgcaataatatttccaagaatttaggggttttcaagaagtatgatcttcaatca 2665  
QY 1264 ttaagcatgagcagaagaatttattccatgatattcgtcatcttgggtgtatttgcctt 1323  
DB 2666 aaacgcagagaatagtaactgttccacatgttctcagcccttgaattatcttctt 2725  
QY 1324 gtcagttattcagagatatttttaacggcagagcaggggaatttaacgataga 1383  
DB 2726 attacaatttccctcaaggtttccacatttgcagaagctggagagatcctccacaagcg 2785  
QY 1384 ttaagacacttggccttcaaaagccatgttatatacagabattgctgtgttgaagaag 1443  
DB 2786 ctcgatatcatgttttccgatccatgctcagacagagatgtgattgtgttgatccct 2845  
QY 1444 gaaaacagcagagaggttgcacaacataatagccatagatataagcacaatcaagga 1503  
DB 2846 aaaaacacacactgagacattgtactacaggtcgcgcaatgagtcgctccaagttaa 2905  
QY 1504 gcaaacgttccagattgcttcaacacaaatgcaacaaatggaattccaatt 1563  
DB 2906 gctatagttccaggtcgtgttaattacccagaatataagcaaatccttgggacaagata 2965  
QY 1564 alcatcttccttatataatgagtggagatgacatcttcgtatcttgagtattccagta 1623  
DB 2966 attatacttcattcatatgtgtgacactaacacatgtaactcttagcaatttaccatc 3025  
QY 1624 ctgtccgtgacaggaatgattgaacccgacgaatgactggtttgccaacaagaatag 1683  
DB 3026 attgtaaatagcaggagttgttgaatgaanaatgttgcctgcgacaagcactgaaga 3085  
QY 1684 caagaactaaagcagtcggaaagatagcaactggaagcttggagaataatagtaacta 1743

Db	3086	aaagaactagaaagtgctgttggaagatcgtcacttgaaagcaatgaaacttcgcgaacggt	3145
OY	1744	gtgtcattacaagaagaaaacgcttcgcgagcaaatgtatgaaagatgcttcgaactca	1803
Db	3146	gtttctttgtcacaaggcgcaagattgttgacaatactatgtccagaagatttcgaatgcca	3205
OY	1804	cacgaataatccctcgaaagaagacagcagattatggaagctgtttagcattgaacgattgcc	1863
Db	3206	tacagaatacctttgaaagaaagacacactcttggaaattacatttcccttcaacccaagca	3265
OY	1864	ttataatttttgcctatgacgcaagggtttcgaatttggagccattttatccaagcttga	1923
Db	3266	atgagtatttttccctactagctgtgaaagtgttcgcggttgagcctacttggtggcaataaa	3325
OY	1924	cgaatgcgcccaaggagcgatgttcataagtttttactgtcaattgtcatatgagcattagcc	1983
Db	3326	ctctagtcgcttcttgaaagatgtctctgttagattttccagctgtgttcctttgtgtgcattgccc	3385
OY	1984	atcgaaaaaacgcgtcgttttgcttccttgaaattttccaaaagccaataatcggggctgycgat	2043
Db	3386	gtggggcgaagtcagttcaattcgtctcctgtcactgacaaagccaataatccagccac	3445
OY	2044	ctgtttgcctctgttgaaagaagaaacaaatataagacaagccgacgacgaagaggaanaag	2103
Db	3446	atcatcatgcatcatggaanaaaacccctcttgatgtaacagctaaagaaacgaagagcctaag	3505
OY	2104	ccagacacatgttgaaaggaaatttagagtttcgaaagtcctcttctctatccatcgtgc	2163
Db	3506	ccgaacacatttgaaagaaatgtcacatctgtgaaagtgttatccaacttccaccgga	3565
OY	2164	ccgaatgttttcaactcctcgttgctatccctccagatattgaaacgaggaagaaacagtagca	2223
Db	3566	ccggaacatcccaactgcgtctcgaagccgagaccttgaaagtgaagaaaggccagacgcttgct	3625
OY	2224	ttctgttggaagcagcagcgtctcttggaagaagcaactctgttcacactcttcgaagacttat	2283
Db	3626	ctgtgtgttgcaagcagtgctctgttggaagagacagatgcttcacgctccctggaagcgtttcac	3685
OY	2284	gaaccgcgtgcagaagacaagtgtcgttttgaatgvtgtgaaatgcbaaaagaaattgaatgtagag	2343
Db	3686	gaaccctcttggaagaaagtgtcgtcttggatttggaagaataataacgcagctgaattgcag	3745
OY	2344	tgtgtccggtttccaaatagaacactgtctctccaagagcctcgtcttcaacttcgaagcatt	2403
Db	3746	tgtgtccggaagacaacccgtggacatcgtgtccccagagaccacactcgtttgtacgtgcagcatt	3805
OY	2404	gcttgaagaacatcgcctctatgttgacaacacagcgtgtgtgttgcattagatgaaatcaagaa	2463
Db	3806	gctgagaacactgtcctatgtgagaacaacagcagcgtgtgtgtctacagaagaagatcgtgagag	3865
OY	2464	gctgcacaatgcagcaaatatccattctttatbtgaaggtctccctcgtgaataacacaaca	2523
Db	3866	gcacgaagaagagcgacaacatactgtccttcatcgtgtcaactcctaataatataagcatt	3925
OY	2524	caagtgtgactcgaagaagagacacgcttctctgcgcgcagaaaacaaagctatgatttgcga	2583
Db	3926	aaagttaagaaacaagaagaaatctcagctctctgtgtgcagaaacaacagcattgtgcatagct	3985
OY	2584	agggtctctctccaaaacccaataatttatgttgatggaagcgaactgcagcccttgat	2643
Db	3986	cgtgtccctgtttagacgcgtctatatttgcgtctttgaaatgaagccagcgtccagctcttgat	4045
OY	2644	aatgacagtgagaaggtgtgtctacgcatcgtcccttgataaagccagagcaggaagagcaatgc	2703
Db	4046	acagaagaatgaaaaaggtgttcacagaagacgcctcggaataaagccagagaagagccgcacctgc	4105
OY	2704	ctagtgatcactcaacaggtctctgcgaattcagaacgcgcagatttgatagtggttcgcac	2763
Db	4106	atctgttatgtctcacccgctgtccacacatccagaatgcagacttaagtgtgttcag	4165
OY	2764	aattgaaagataaaggaaagaagaactcatcaagcctctccgtgagaatcgcagacatatat	2823
Db	4166	aattgcagaaatgcaaggagcatctggacaagcatctcagcagctgtctgtgcacaaagaagctctat	4225

QY	2824	t	t	a	a	g	t	a	g	t	a	a	t	g	c	a	a	c	a	g	t	c	a	g	2848
Db	4226	t	t	t	c	a	t	g	t	c	a	g	t	g	t	c	c	a	g	t	c	a	g	4250	

RESULT	5
AAZ94738	
ID	AAZ94738 standard; cDNA; 4646 BP.

AC AAZ94738;

DT 01-AUG-2000 (first entry)

Human ATP binding cassette ABCB1 (MDR1) cDNA.

KW ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;

lupus erythematosus; diagnosis; gene therapy; MDR1;

XX

XX

XX

[illegible]XX  
XX

XX

	X
2	
L	
T.	
+.	
I	
D	
S	
E	
A	
H	

XX  
0000000000000000

XX  
XX

PT for treating atherosclerosis and other inflammatory disorders -

PS Claim 9; Page 110-112; 154pp; English.

CC The present sequence is that of human ATP binding cassette  
CC subfamily B protein ABCB1 cDNA. The cDNA was identified in

CC differential display method in which monocytes from peripheral blood were subjected to macrophage differentiation and cholesterol

CC loading with acetylated low density lipoproteins and subsequent

CC cholesterol sensitive genes. The gene maps to chromosome 7q21  
CC and is also termed MDR1 (multidrug resistance). The invention

CC provides cholesterol-sensitive ABC genes (see AA294734-63). The  
CC genes and polypeptides encoded by them can be used for diagnosis

CC and therapeutic applications, and for biochemical or cell-based assays to screen for pharmacologically active modulator compounds

CC useful for the treatment of lipid disorders, atherosclerosis or  
CC other inflammatory diseases such as psoriasis and lupus

CC erythematosis.  
XY

Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Query Match 36 18: Score 1030 2: DB 21: Length 4646:

Best Local Similarity 62.28; Pred. No. 2.6e-263;  
Matches 1695; Conservative 0; Mismatches 988;  
Indels 42.

[illegible][illegible]

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

[illegible]

OY	286	aagatcttgaaaagtcgtcgaatctccgaataaagatctcgtgagagacagtcgcttgctgtcgt	345
Db	1646	aagatcttgaaaagtcgtcgaatctccgaataaagatctcgtgagagacagtcgcttgctgtcgt	1705
OY	346	ctcaatgacagcttgagaaagatcagtgatgtcccaagctcttcacagaggtatcatgataccgcat	405
Db	1706	aacagctgacgctgagaaagatcagtgatgtcccaagctcttcacagaggtatcatgataccgcat	1765
OY	406	gatgctcttcaatctgctgagaaatgacatcagaagctcttcaatgctgagcatctacga	465
Db	1766	gagggagatgctcagtgctgtgctgagaaatgacatcagaagctcttcaatgctgagcatctacga	1822
OY	466	gaaccatctggaatgctgagaaatgacatcagaagctcttcaatgctgagcatctacga	525
Db	1826	gaatcatctgctgagaaatgacatcagaagctcttcaatgctgagcatctacga	1885
OY	526	atcaagatgagagagatgctgacatgagaaatgagagagacagcagaagagagaa	585
Db	1886	atcgcacatgagagagatgctgacatgagaaatgagagagacagcagaagagagaa	1945
OY	586	aatgcataatcttcaacagagatcttcaataaattcaatacatctgttgagagaaa	645
Db	1946	aatgcataatcttcaacagagatcttcaataaattcaatacatctgttgagagaaa	2005
OY	646	gagagctcaaatgagtgagagagcagaacagagagatgcagatgctgctgactatctga	705
Db	2006	gagagagctcaaatgagtgagagagcagaacagagagatgcagatgctgctgactatctga	2065
OY	706	aaccccaagatctcgaatttaagatgagagctcagctcgccttgatctcaaaaagcagatca	765
Db	2066	aaccccaagatctcgaatttaagatgagagctcagctcgccttgatctcaaaaagcagagca	2125
OY	766	gctctcaagctcagctcgtgagagagcagagacagagctcagctcagctcagctcagctcag	825
Db	2126	gctctcaagctcagctcgtgagagagcagagacagagctcagctcagctcagctcagctcag	2185
OY	826	cgaactctactactcgaaagtcagagattgatgtgacacctaaagagatgaaatgctgcgcg	885
Db	2186	cgtcttgctacagctcgaattgatgtgacacctaaagagatgaaatgctgcgcg	2245
OY	886	gagagaaagagacaaatgctgacactaaatgagcacaacagcggtctaatatctatccacttgatg	945
Db	2246	gagagaaagagacaaatgctgacactaaatgagcacaacagcggtctaatatctatccacttgatg	2305
OY	946	tcaagacgatattaaaagcctgatacagatgagatgcatactgacatctca-----	997
Db	2306	atgcagacagcagaaatgagatggaattgaaatgagcagcgtgatctccaaaagtga	2366
OY	998	-----ctgaaagaaagacacaactcactctctgcacctctgagaaagacatca---	1045
Db	2366	atgatgccttggaatgctctcacaatgatctcaagctccagctcagcttaataagaaagatca	2422
OY	1046	-----agtcagacttattgacaaagcgctgaggaatcccaactat	1088
Db	2426	actcgtatgagatgctcgtgatacacaagcccaagacagaagaagctltgataccaaagagct	2485
OY	1087	aaagagataagctctcgtgagatctcctcttcaaaaattttaaagtatacaagcgttga	1146
Db	2486	ctgagatgaagataactccagcttcccttlttgagagatatagaaagtataaatttaactctga	2544
OY	1147	tgcgcttctgtgctcgtgagagacatgctctcgtctcaatgaaactgctatccagta	1206
Db	2546	tgcgcttctgtgctcgtgagagacatgctctcgtctcaatgaaactgctatccagta	2606
OY	1207	tcttcacatcactcttgcaaaaatatacacaagtcttgagaaataatgataaaac--aca	1266
Db	2606	tcttcacatcactcttcaaaagatataagaggtcttcaagaatltatgatacctgaaaca	2666
OY	1264	ttaagcagctgagaaatattatcaatgataatctgcatacttgcttggttggttctt	1322
Db	2666	aaacagacagaatagtaactgcttctcaatgcttctcctagccttgagaaatattcttct	2722
OY	1324	gtcagttattctcagcagggatattttaaagcagagcagggagaaattttaaagctgaga	1383

Db	2786	atcaccatttccctccaggggttcacattggaagaagccggagagatccccaagcgg	2785
Qy	1384	ttaaagacccctggcctccaagaccgattatcaagatattgccttggtttgatgaaag	1443
Db	2786	ctccgatacaatggtttcccgataccagctccagacagagatgtagttgttttgagacct	2845
Qy	1444	gaaacagcacaggaaggcttgacaacaataatgacatagatatagacacaattccaag	1503
Db	2846	aaaaacaccacttggagattgacttaccagagctcgcgaatgatcgtcccaagttaaagg	2905
Qy	1504	gcaacaggttccaggtatggtcgtcttaacacaatagcaactcaacatgaggatttcagt	1553
Db	2906	gctataggtttccaggtctgtgttaattaccaggaatataagacaactcttggacagaa	2965
Qy	1564	atcatttccttatataatgatgatggagagagacattccgtattcttgatatactgcagt	1623
Db	2966	attatacccttcaactctatggttggcaactaacaactgttactctcttggcaattgtacc	3025
Qy	1624	cttgcctgtaacagaaatgatctgaacccgcagacaatgactgtgatttggacaacaaga	1683
Db	3026	attgcaataagaggaagtgtgttgaaatgaaatgtgttgcgacagacgtgaagaatag	3085
Qy	1684	caagaacttaagcatctgtgaaagaatagacaactgtgaagcttggagaaatatacgtact	1743
Db	3086	aaagaactagaagggtgtctggagaagatcgtactatgaaagcaatagaaacttcggaa	3145
Qy	1744	gtgtcatttaaaagggaaagagccttcgagcaaatatagaagagatgtcttcagacct	1803
Db	3146	gtctctcttgatccaggagcagaaagtctgaacataatgatacgtccagagtttgagatga	3205
Qy	1804	cacagaaatacctcgaaagaaagcacagatattggaagctgttaatgatacttcagacat	1863
Db	3206	tacagaaactctcttgaggagaagacacacatctcttggaatttaacttctccaccag	3265
Qy	1864	ttatataatttgcctatgcagcaggggttcgaatttggagcctatataactgaagctg	1923
Db	3266	atgatagtatttctcctatgctgtaatgtttccgggttttggagcctactgtgtgacaata	3325
Qy	1924	cgaagaccctccagaggggcagatgtaatgattttacgcagatctgacatagagatctg	1983
Db	3326	cccaagagcctttggagagatgtctcgttaagtaatttccagcgtgtgtcttggvgcag	3385
Qy	1984	atcggaaaaagccgctgttttgcgtccctgaatacttccaagaccaaatcgggggtctg	2043
Db	3386	gtggggcgaaagctatgcatttgccttcctgcgtactatgccaagaacccaatacagc	3445
Qy	2044	ctgtttgcctgttggaaaagaaaccaaatatagacagccggatctcaagaaggaanaag	2103
Db	3446	atcatalcatcatcttgaaaacaccccttgaatgacagatcagacaggaagccttaag	3505
Qy	2104	ccagacacatgttgaaaggaattttagagtttcggaaagctctcttcttcatacagctg	2163
Db	3506	ccgaacacacatctggaaggaatgaatgaacatcttggatgtatcttcaactatccacc	3565
Qy	2164	ccaagatgttttccatccctcgtgtgtactacccctcagatattgacggaagaaagca	2223
Db	3566	ccggaacatcccaagcttccagagacatgagccttggaagtgaaagaaggccagacg	3625
Qy	2224	tttgttgggagcagcggcgtgtgggaaagacacttcttcaactcttcgacagacttat	2283
Db	3626	ctgtgtgggcagcagctgtgctgtgggaagagacacagtgctccagctccctggag	3685
Qy	2284	gacccctgtgcaaggaacagatgctgttgatgagtgtgtgatacgcaaaagaattga	2343
Db	3686	gaccctcttggcaggaagaaatgctgtctgtgaatggcaaaataaagcagatgaattc	3745
Qy	2344	tgagctccgttcccaataagacatggttcttcccaagagcctgtctcttcaacatg	2403
Db	3746	tgtgtccgagcaacccctgggcatcgtgtcccaaggagccaatcgtttgactcgagcat	3805
Qy	2404	gctggaacatcgctatgtgtgacaacagccgtgtgtgtccattagatgagatcaagaa	2463

Db	3806	gctcgaacatctgcctatctgagacaacagccggctgctgctcagtgaaagatcgtgag	3865
QY	2464	gcccacaatgcagcaaatatccattcttttatgaaagctctccctgagaaatatacacaa	2523
Db	3866	gcgcgaagaagagagccaaatcaatcactgcttcctcatcgctacctctctaataataatagca	3925
OY	2524	caagttctgactgaaagagagacacagctttctctgcgcgcagaaacaaagctctgctattgca	2583
Db	3926	aaagttagaagacaacaagaaatcagctctctctgtgagcaaaacaacagcatctgcatagct	3985
OY	2584	agggctctctcccaaaaaccacaattcttatgttgtagagagcaactccagccctgat	2643
Db	3986	cgctcccttgtagaacagctctcatatttgccttttgtagaagcgaagctgcctctgat	4045
OY	2644	aatgcacgtggaaggtgtgttcacgcatgcccctgataaaagccagagacgggaagacatgc	2703
Db	4046	acagaagaatgaaaggtgtgtccaaagaagccctgcgaacaagacagaagaagccgcacctgc	4105
OY	2704	ctagtgtcactccacagctctctgcgaatctgaagacagatttgatggtgttcctgac	2763
Db	4106	atctgtgatgtgtccgcctccctccacacatccagaatgcagacttaatctgtgtttcag	4165
OY	2764	aatgcagaagactaaagagacaaagaaactcatcaagagctctctgagaaatcgagatacat	2823
Db	4166	aatgcacagatcgaagagagcatgtgcacagcatcagctgctgtgcacaagaagacatcat	4225
OY	2824	tttaagttagtgaatgcacagctcag	2848
Db	4226	ttttcaatgtcaggtgtccagctgcg	4250

XX	RESULT	6
XX	AAAN70752	
XX	ID	AAAN70752 standard; cDNA; 4669 BP.
XX	AC	AAAN70752;
XX	DT	21-MAY-1991 (first entry)
XX	DE	Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones
XX	XX	lambda-HDR10.5 and 104.
XX	KM	Chemo-therapy resistant tumour cell; P-glycoprotein; ss.
XX	OS	Homo sapiens.
XX	FH	Key
FT	FT	Location/Qualifiers
FT	CDS	425..4267
		/*tag= a
XX	PN	W08705943-A.
XX	PD	08-OCT-1987.
XX	PF	26-MAR-1987; 87MO-US00758.
PR	PR	01-AUG-1986; 86US-0892575.
XX	PR	28-MAR-1986; 86US-0845610.
XX	PA	(UNIT ) UNIV OF ILLINOIS.
XX	PI	Robinson IB, Pastan IH, Gottesman MM;
XX	DR	WPI; 1987-291656/41.
XX	XX	P-PSDB; AAP70452.
PT	PT	DNA for multi-drug resistance in human cells - used to detect
PT	PT	chemotherapy-resistant tumour cells and for producing
XX	XX	polypeptide(s) for diagnosis and therapy
XX	XX	Claim 2(a); Table 5, pp30-39; 61pp; English.
XX	XX	The human multi-drug resistant KB carcinoma cell lines were used as

CC the source of the mdrl gene nucleic acid sequences (AA070751). To  
CC obtain cDNA clones of the mdrl gene (AA070752), poly (A) and RNA was  
CC used. Analysis of the Aa sequence presented in (AA070752) indicates  
CC that the mdrl gene product is likely to be a transmembrane protein.  
CC The presence of transmembrane domains and potential glycosylation  
CC sites is consistent with the mdrl protein being related to the  
CC P-glycoprotein.

SQ Sequence 4669 BP; 1393 A; 896 C; 1126 G; 1254 T; 0 other;

Query Match	36.1%;	Score 1030.2;	DB 8;	Length 4669;
Best Local Similarity	62.2%;	Pred. No. 2.6e-253;		
Matches 1695; Conservative	0;	Mismatches 988;	Indels 42;	Gaps 3

QY	166	atattgtataagaataccacgatatagataacttttccacgcgctgatatataactgtatcc	225
Db	1526	ataattgtataataagacaagatattgtacagctatctcgaagaagtgtggacaaacacgaataat	1585
OY	226	atagaagagacgtgtgaattataaaatgtttctttcaattatccatccaagaacatatac	285
Db	1586	attaagagggaatttggaaattccgaataatgttcaactcttcagttaccatctcgaanaagaagt	1645
OY	286	aagattctgaaagcttcgaatctcagaatlaagtctcgagagacagtgccttgcgtgcgt	345
Db	1646	aagattcttgaaggccttcgaaccctgaaagctgcagaagtgtggacagacggtgtgcctgtttga	1705
OY	346	ctcaatgtgacgttgaggagaaggtgaagctgagtcacagctcttcgaaggtttatatactcgat	405
Db	1706	aaacagtgcgcgttgaggagaagcacaacacagctccagctgtatgacagsgctctatgccccaca	1765
OY	406	gacgtgcttatcatctgtgtgatgagaatgatcacatcagsgctttaaagtgcggcatatcga	465
Db	1766	gagggagatgtgtcagctgttgatgacacgagatatataggacataaattgaaggtttcttcagg	1825
OY	466	gaaccatattggaattgtttgttgaaggcctgtttttcttgaggccaccatcagtlaacaat	525
Db	1826	gaatcatattgtgtgtgtgacgtgcgaaccctgtattgttgcaccaccgatagtgtgaatc	1885
OY	526	atcaagtatgaaagagatgtatgtacatgataagaagaatgtgagagacagacaagggaga	585
Db	1886	attcgtctatagccgttgaaataatgtccacatgtatgagatgagaatgacgtgcagaagacc	1945
OY	586	aatgctgatgatttatcatctgacgttctcctaataattatcatattgtgtaaggggaaaa	645
Db	1946	aatgctctatgactttatcatgaaactgtccctcataaattttacacccctgtgttggagaga	2005
OY	646	ggagctcaaaatgtgctggaaggcagaaacaaagatctgcaattctcgtgtcgttaagtcca	705
Db	2006	ggggccagttggtgtgtgtggcagaagcagaatgatctgcatttgcacgttgcctgtgtgc	2065
OY	706	aaccoccaaagatcttgattttagatgaggtctacgtctgcgcttgatctcagaagaagaatca	765
Db	2066	aaccoccaaagatctctctgcgtgcgtgtgagccagcttcaagctcttggacaagaagaacgaagca	2125
OY	766	gctgttcaaaactgtcaactgtgaggaagcagagaagaagctctgcagctcaatctgcgttatgacac	825
Db	2126	gtgtgttcaagttgctcctcgtgataaggccagaagaagaagttctggccacacatgtgtatagcccat	2185
OY	826	cgacttctactatctcgaagtgcagatttgattgttgacccttaagaagatgaaagtctgcgt	885
Db	2186	cgttgtctacagtttgcgttaagtctgcagctgcatcgtgtgtcttgatgattgtgaatcatgtgtg	2245
OY	886	gagaagaagacacatgctctgaactaatgtgcgaagaagaagctctatatattcaactgtgatg	945
Db	2246	gagagaagaagaaatcatgtatgaactcaatgaagaagagaagcaattactccaactgttcaca	2305
OY	946	tcaacagatattaaanaagctgatgaaacagatgagctcaatgacatatattca-----	997
Db	2306	atgcagacagcaggaatgtgaatgtgaatttgaanaatgcagctgtatgatacccaagaatgaa	2365
OY	998	-----ctgaagaagaaccaaactcaacttctcctctgcactctgtgaagaagctga---	1045

Db 2366 atgatgcttggaaatgtcttcaaatgtatccagatccagttcataagaanaaatgta 2425  
Oy 1046 -----agtcagacttcaattgcaagcgtgaggaatcccaactct 1086  
Db 2426 actcgttagagtgctcgttgatcacaaagcccaagacagaagctagtaaccaagaagcct 2485  
Oy 1087 aagaagataagctctcgtgaagctctctatataaanaatttaaaagttaaacaagctgaa 1146  
Db 2486 ctgagatgaagatatacccccagttcttcttggagagatlaagaagctaaatlaactgaa 2545  
Oy 1147 tgcgccttggctcgtggagacattgctctgttctaaatbgaaactgttcatcagta 1206  
Db 2546 tgcgcctatttgtgtgtgtgtatcttggccattataaagtggcgctgaacacga 2605  
Oy 1207 ttcttcacatccttgcgaanaattataacacatglttggaaataatgataaaac---aca 1263  
Db 2606 ttgcaataaattcttcaagaattalaagggttlttacaagaatgagatcctgtaaca 2665  
Oy 1264 ttaagcagatgtagcaagaattatccatgatactgcatttgggtgtattgtctt 1323  
Db 2666 aaagcagcaagatagtaactgttcttcaactatgttctcagccctggaaatatttctt 2725  
Oy 1324 gtcaattatctacgacaggaattalttacygacagcagcagggaaatttlaacgatlga 1383  
Db 2726 attacatttcccttcaggttccaacttgcgaagcgtggagagatactcccaacgagc 2785  
Oy 1384 ttaagacacttgccttcaagccatgttatactagaatattgctgttggatgaag 1443  
Db 2786 ctccgatacatggttcttcgataccatgctcagacagagatgtgagttgttltatgacct 2845  
Oy 1444 gaaacacacagcagcaggttgcacaacaatatatagccatagatagacaaattcaaga 1503  
Db 2846 aaaaacacacttgagcacttgaactccagctcgcgaatgatagtcgcgaagttaaagg 2905  
Oy 1504 gcaacaggttccagagattgagcgtcttaacacacaaatgcaactaacaatggacttca 1563  
Db 2906 gctatagattccagcgttgcgttaattaccagaatalagcaaatcttggcagagaata 2965  
Oy 1564 atcatttcttataatagatggagatgacatccctgattctgaatgttctccagta 1623  
Db 2966 attatacttcatctatggttggcaactaacaactgttactcttgcatttgcattccac 3025  
Oy 1624 ctggcgtgtagaattgtaaaacgcgcaatgactggaatttggccaacaagaattag 1683  
Db 3026 attgcaatagcagaggtgtgtgaatgaatgaatgttcttgcgaagcactgaagaatag 3085  
Oy 1684 caagaacttaagcattgctggaagaatagcaactgaagcttggagaaatalactata 1743  
Db 3086 aaagacactagaagtgctggaagaatcgctactgaagcaatagaacttccgaacgctt 3145  
Oy 1744 gttcattaacaagaagaaagccttcgacgaactgtatgaagagatgcttcaactcaa 1803  
Db 3146 gttctttagactcagagacagaagatttgaacatagtatgctcagagatttgcaggt 3205  
Oy 1804 cacagaatactcgtgaagaagaacagattatggaagctgtatgacttccatccatg 1863  
Db 3206 tacagaactccttggagaaagcacacacttggaaatlaacatttccctcccaaggca 3265  
Oy 1864 ttatataatttgcctacgacagcaggttctcgatttggagcctatttlaactaaactg 1923  
Db 3266 atgataatttcttccatgctgagatgttccggttggagcctacttggtagcaataaa 3325  
Oy 1924 cgaatgaccccaagagcagatgttcatagtttctactgaatgcatatgagactatg 1983  
Db 3326 ctcaagaactttagagagatgttctgtatatttcaagcgtgttcttggtagcctg 3385  
Oy 1984 atcggaaaaaacgctgcttgcctcgaatatttcaagaagccaactcggggctcgat 2043  
Db 3386 gtggggcgaagtcatttgcctcgaactatgccaagccaataatatacgaagccccc 3445  
Oy 2044 ctgttgccttgttggaaaaaacaataatagacagccgcagtcacaagaaggaagaa 2103  
Db 3446 atcaatgatactatgaaaaccccttgtatgacagctaaagcaggaagcctaagt 3505

Oy 2104 ccagacacatgtgaaagggaatttagagtttcgagaagctcttctctatccatgctgc 2163  
Db 3506 ccgaacacatgtgaaagaaatgacacattgtgtgaagtgtatattcaactacccaaccca 3565  
Oy 2164 ccagatgttctatccctcgtgtgtctatccctcaaglatatgagcgaagaagacagatga 2223  
Db 3566 ccgagacatccagtgcttccaggaactgagcctggaggggaagaagggccagacgttgc 3625  
Oy 2224 ttgttgggagcagcgtgttgggaagaagcaactctgttcaactctgcagaagcttatt 2283  
Db 3626 ctgttggcagcagtggtgtgtggaagagcacagtgttccagctctctgtgagcgttctac 3685  
Oy 2284 gaccccgagcagcagcagtgctgttgaatgtgtgagatgcaagaagatttgaatga 2343  
Db 3686 gaccccttggcaggaagatgctgtctgtatgtaacaaagaataaagcagatgatagtc 3745  
Oy 2344 tgcctcgttcccaatatgacatcgtctcctaagaagcctgtctcttcaactgcagcat 2403  
Db 3746 tgcctcgaagcacacacttggcatcgtgtcccaagggcccaactcgttgtgactgcagcat 3805  
Oy 2404 gctggaacatgcgctatgttgaacaacagcgtgtgtgtccattagatgatacaagaa 2463  
Db 3806 gctggaacatgtcctatgtgagacacacagccggtgtgttcaagaagaagatcgtgag 3865  
Oy 2464 gccgcaaatgcaagcaaatatccattcttatttgaaggtcttccctgagaataataca 2523  
Db 3866 gcaagaagaaggccaacatatacttcttcaatcgaatcagctgtcctaataataataga 3925  
Oy 2524 caagtgtgacttgaaggagcagcagcttctgtgcgcgaagaacaagaactagattca 2583  
Db 3926 aaagttagagaaagaaggaactcgtctcctgtgtgcgaagaacaacagcattcgaag 3985  
Oy 2584 aggcctcttcccaaaaacccaatttatttatttgaatgagccactcagcctcgat 2643  
Db 3986 cgtgccttgttagacagcctcatatttctgttcttgaatgaagccaagctcagcctg 4045  
Oy 2644 aatgacgttgaagaagtggttgcagatcccttgaatgaagccaggaacgggaagcag 2703  
Db 4046 acagaagttgaagaagttgttccagaagccctggcaagaagccaggaagccgcacgtgc 4105  
Oy 2704 ctgagtgcactcacagcgtctcttcaattcagaacgcagatttgaatgttctgcac 2763  
Db 4106 attgtgattgtccacgcgtctgtcccaatccagaatgcaacttaatgtgttctcag 4165  
Oy 2764 aatggaagataaaggaaacaaaggaactcatalcaagagctcttgaagaatcgaacata 2823  
Db 4166 aatgacagatcaaggaagcatgtgcacgcatcagcagctgtgtgcacagaagacatcat 4225  
Oy 2824 tttaagttagtgaaatgcacagtcag 2848  
Db 4226 ttltcaatgtcagtgctccaggtcg 4250

RESULT 7  
AA113394  
ID AA113394 standard; DNA; 6505 BP.  
XX  
AC AA113394;  
XX  
DT 24-JUN-1996 (first entry)  
XX  
DE Hybrid vector pSF-MDR.  
XX  
XX Hybrid; vector: gene transfer; gene therapy; haematopoietic stem cell;  
KW retroviral; murine embryonic stem cell virus; MEV5;  
KW Mooney murine sarcoma virus; (MoMuSV);  
KW Friend murine leukaemia virus; F-MuLV; ds.  
XX  
OS Synthetic.  
XX  
PN DE19503952-A1.  
XX





D	b	4538	gttcctttgactcaagagcaggaagtttgcacatagtatgctccagagtttgcaggtaccca	4597
Q	y	1804	cacagaaataccctcgaaagaagacagatatttggaaagcgtgttatgcattcaagcatgc	18639
D	b	4598	tacagaaactctttggagaaagacacatctcttggaattacatttctctcaaccagga	4657
Q	y	1864	tttatatatttgcctatgcagcagagtttcgatcttgagcctatttaatacgaagtga	19233
D	b	4658	atgagtattttctctcatgctgatgttcttcggtttgcagcctactctgttgcacataaa	47171
Q	y	1924	cgaatggccccaagagcgagttcagatcgaattttactgcgaattgcatacgagctatgcc	19839
D	b	4718	ctcaatgagcttgcagatgtctgttgaattttccaagctgtgtctcttgcgtgcataggc	47777
Q	y	1984	atcgaaaaacgcgtcttgcgtctctgcataattccaagccaatctggggctgtcgat	204339
D	b	4778	gtgggagaaagtcagttcatcttctgtcctgtcgtatgcacaaagccaatatcagcagccac	4837
Q	y	2044	ctgtttgcctctgttgcagaaagaaaccaaatatagacagccgagtcagaaagggaaaaag	21039
D	b	4838	atcctcatgatactgaaaaaaaccccttgcattgcagctacagctacagcaggaagcctaatg	4897
Q	y	2104	ccagagacacatgtbaaagggaatttagagtttcgagaagctctctttctatccatgctgc	21639
D	b	4898	ccgaacacatctgaaaggaatgtccacatcttgcgtgaagattgtatccaactttccaccga	4957
Q	y	2164	ccgaattttcttccatccctgcgtctatccctccagatattgaacgaggaagaaacagtatga	222339
D	b	4958	ccggacactcccaatgctcttccaggaactcgagctcggaggtgaagaaagggccagacgtgtgc	5017
Q	y	2224	cttgcgtggagcagcgctgttgcgaaagcaactctcgttcaactctcgcagagacittat	22839
D	b	5018	ctggctggagcagcagctgtcgttgcgaaagacagacgtgtccagctccctcgagcggtctac	50777
Q	y	2284	gaccccgtgcagaagacacaagcgctgttgcattgtgtgcatagcacaaagaattgaattgaag	234339
D	b	5078	gaaccccttgcagaggaagagtgctgtcttgcattgcgcacaaataaacagactgaatttcaag	51377
Q	y	2344	tgcgtccgtctcccaataagacatcgtctctccacgaagcgtgtccttcaactgcagacatt	24039
D	b	5138	tgcgtccgagacacactcctgggcattcgtgtcccaaggagcccatccctgttgcactgcagcatt	5197
Q	y	2404	gctgcagaacatcgcctcattgtgtgaacaacagccgtgtgtgcattagatgaagataccaagaa	24639
D	b	5198	gctgcagaacattgcctatgtgagacaacagccgggtgtgtcaccaggaagagtgctgtgag	5257
Q	y	2464	gcgcgcaaatgcagcaaatatcatctctttatttgaagctctccctgcagaaataacaaca	25239
D	b	5258	gcacaaaagagagccacaatatactgtcctatcagatctcctcaataatataagcatt	53177
Q	y	2524	caagtctgcactgaaagagacagacgtcttctgcgcgcagcaaaacaagcctagcatttga	25839
D	b	5318	aaagtgaagagacaaaaggaactcagctcctctgtgtgcacaaacaacagcatgtgcataagc	53777
Q	y	2584	agggtctctctccaaaaaacccaataatttattgttgcataagagcacttcaagccctcgat	26439
D	b	5378	cggtgcctctgttagacagcctcatatttgccttttgatgtgaagcagcagctcagctcgat	5437
Q	y	2644	aatgcacgtgagaaagtggtgtctgcagatgcctctgtataaagccaagacgggaagacatgc	27039
D	b	5438	acaaagaaggttgaaaaggtgttgcacaagaagcccttgcacaaagacagagaaagccgcacctgc	5497
Q	y	2704	ctagtgtgcactcaacagcgtcctctgcgaatttcagaaacgcagatttgaatgtgtctgac	276
D	b	5498	attgtgattctcaccgctccgtccacacccacccaagaaatgcagacttataatgtgttccag	5557
Q	y	2764	aatgcagaagaataaaggaacaagaaacttcaacatcaagatcctcttgcagaaatcagataatat	28239
D	b	5558	aatgcagagagtcaagggagcattgcacgacatcagcagctgtctgcacagaagaagcattcat	5617
Q	y	2824	tttaagttaatgaatgcacagttcaag	2888
D	b	5618	ttttcaatgtgtcagttgcacgctgcgtg	5642

RESULT	8
AAZ49332	standard; CDNA; 3860 BP.
AAZ49332	
AAZ49332;	
14-MAR-2000	(first entry)
Human wild-type multidrug resistance-1 (MDR-1) cDNA.	
Multidrug resistance; MDR-1; P-glycoprotein;	
transmembrane efflux pump; haematopoietic stem cell; transduction;	
bone marrow transplantation; chemotherapy; radiation therapy; cancer;	
gene therapy; gene replacement; genetic defect; thalassemia;	
Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;	
cytokine; wild-type; ds.	
Homo sapiens.	
Key	Location/Qualifiers
CDS	1..3843
	/*tag= a
	/product= "Human wild-type MDR-1 protein"
mutation	replace (553..555, GTT)
	/*tag= b
	/note= "CDNA sequence of G185V human mutant MDR-1 given in AAZ49333"
WO9961589-A2.	
02-DEC-1999.	
27-MAY-1999;	99WO-US11825.
28-MAY-1998;	98US-0086988.
(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.	
Sorrentino B, Bunting K;	
WPI. 2000-072615/06.	
P-PSDB; AAF58186.	
Ex vivo expansion of hematopoietic stem cells transduced with a	
sequence encoding human multidrug resistance-1, used for bone marrow	
transplantation -	
Claim 10; Page 68-70; 113pp; English.	
This sequence represents cDNA encoding human wild-type	
multidrug resistance protein MDR-1. MDR-1 is a transmembrane	
efflux pump, responsible for the export of drugs from cells,	
particularly cancer cells. Wild-type MDR-1 shows increased	
resistance to etoposide and decreased resistance to vinca	
alkaloids compared with a mutant form (AA158187) where the Gly at	
position 185 is replaced by Val. The invention relates to transducing	
haematopoietic stem cells with nucleic acid encoding an MDR protein	
and culturing the modified cells. The modified haematopoietic stem	
cells are useful in bone marrow transplantation (to reconstitute	
haematopoietic systems in patients who have undergone chemotherapy or	
radiation therapy) and in ex vivo gene therapy of genetic defects in	
cells derived from haematopoietic stem cells, e.g., thalassemia,	
Gaucher's disease, sickle cell anaemia or leukaemia. The modified	
cells can also be used to identify factors involved in regulating	
proliferation and differentiation in haematopoietic stem cells.	
Haematopoietic stem cells that express MDR-1 will be protected against	
chemotherapeutic agents, so can be grafted while the patient is	
undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells	
provides sufficient cells to permit standard biochemical analysis.	
Overexpression of MDR-1 allows cytokine-driven expansion of	
haematopoietic stem cells by at least 10-fold compared with a maximum	







Db 1402 gaataccttggtggaatcaggaacccgtgattgtttccaccagatagctggaaac 1461  
QY 526 atcaagatgacagagatggtgactgatatgaagatgagagagacgaagggaagca 585  
Db 1462 atccgatagtccgtgaaatgacccaatgagatgagatgagaaactgccaaggagcc 1511  
QY 586 aatgcgatagtattacatagaggttccctaataaattatacatgtgtaggggaaaaa 645  
Db 1522 aatgcatagtacttatcatcagaactgcctcataaatttgacaccccggttgtagagaa 1561  
QY 646 ggaagctaaatgagtggaaggcagaacacagagatgcgaattgctgctgattcgca 705  
Db 1582 ggggcccagatgagtggtggtgagcagaagagatgcacgtccattgcacgtgcccgttcgc 1641  
QY 706 aaccccaagattcgtgatttagatgagtagcgtcgtccctggttttaagaagaagtca 765  
Db 1642 aaccccaagatccctcgtgtagatgagccacgtccagccttgacacagaagaaggagca 1701  
QY 766 gctgtcaagctgactggaagcgaagcgagcaaatgctgacataacatcgtgtagacac 825  
Db 1702 gtggttcaggtggtcctgtagtaagcagaaaggtcgagaccacatgtgtagatgccat 1761  
QY 826 cgacttctctatctgaaagtgcagattgattgtagaccctaaaggatggaatgctgcg 885  
Db 1762 cgtttgctacatgctglaatgctgcacgtcatcgtgttcgtagtagagatcatgtg 1821  
QY 886 gagaagaagacatgctgactaatgacaaacgaagctcatatattacatcgtgtag 945  
Db 1822 gagaagaagaaatcatgtagtaacatgaaagagaagaagcatcttcacaaactgtgaca 1861  
QY 946 tcaacagatatltaaaaagctgtagacagatgagatcaatgacataatcta----- 997  
Db 1882 atgagagcagcaggaatgagattgaaatagaaatgacgtgtagatccaaaagtgaa 1941  
QY 998 -----ctgaaagaagaacacacacttccctcgtcactcgtgtaagaagctca--- 1045  
Db 1942 atgagtgccttggaatgcttccaatgattcaagatccagctcactaataagaanaagatca 2001  
QY 1046 -----agtcagactcatgacaagcgtgagatccacacacactc 1086  
Db 2002 accgtgtagaggtgcgtgtagatcacaagcccaagacagaaagcttagatccaaagagct 2061  
QY 1087 aaagagataagcttccctgtagactctctatltaaaaatttaaaagttaacaagcctgaa 1146  
Db 2062 ctgagataaagatactccaagttccttcttgtagagattagaaagcttaaatlttaacgaa 2121  
QY 1147 tggccttctgtgtctcgaggacattggtctgttctaataatgaaactgttcatccagta 1206  
Db 2122 tggccttatttgtgtgtgtgtatttggcattataaattgagagcctgcaacacagca 2181  
QY 1207 ttttccatcacttgcgaanaattataacaagtgttgaataatgataaagc-----aca 1263  
Db 2182 ttggcaataatatttcaagaattataggggttttcaagaagttagatgcctgaaaca 2241  
QY 1264 ttaagcagatgcaagaatttattccatgatalcgtcatlttggtgtgtattgtctt 1323  
Db 2242 aaacgcagagatgaaactgttttcaacttgttctagcccttgaattatttcttt 2301  
QY 1324 gtcaagtatttcatgtaggagataattttagcgcagacagagggaattttaagatgaga 1383  
Db 2302 attacatttcttccccaaggttcaacttggcagaagcgtgagagatccctcaacaacgcyg 2361  
QY 1384 ttaagacatttgcctcaagaccatgtlatatcaggaatttgcctgtttagagaag 1443  
Db 2362 ctccgataactgttttcccatccatgctcagacagagattgagttggtttagagaccct 2421  
QY 1444 gaaacacgacagaggctgtgacacaaatltagccaatagatagacaaatltcaagga 1503  
Db 2422 aaaaacacacttgagattgactacacagcgtcgcaatgattgcttcaagttaagggg 2461  
QY 1504 gcaacaggttccaggaattgctgttcaacaaatgcaacaaatgcaacatggaacttccagt 1563  
Db 2482 gctaatggttccaggtctgtgtatccaccagaataatagcaaatcttggagcaggaata 2541  
QY 1564 atcaattcccttatatatgtagtggagatgacattccctgattctgagatgtctccagta 1623  
Db 2542 atataatcccttactatgtgttggcaactaacaactgttactcttagaanaatttagaccatc 2601  
QY 1624 ctggcgtgacaggaatgatttgaaccgcagcaatgacatgagatttggcaacaaagaaag 1683  
Db 2602 attgcaatagcaggaagtgttgaataatgaaatgttgccttgacagacatcgaaagataag 2661  
QY 1684 caagaacttaagcatgctgtgaaagataagcaactgaaagccttggagaataatagctata 1743  
Db 2662 aaagaactaagaaggtgctgtggaagatcgtcatcgaagcaatagaanaacttcgaaacggt 2721  
QY 1744 gtgctatcaagaagggaanaagccttcgagcaaatgtgataagaagatgcttcagaccaa 1803  
Db 2722 gttcttcttgactcagagcaggaagatttgacacataatgataatgctagagattgcaggtacca 2781  
QY 1804 cacagaataactcgaagaagacacagatatttgaagcgtttagatcattcaagccatgccc 1863  
Db 2782 taacgaacactcttgaggaagaagcacacatcttggaaattacatttccctcacaccagga 2841  
QY 1864 ttataatatttgcctatgcagcaggtttcogatttggagccatatttaactaaagctgga 1923  
Db 2842 atgattatatttctctatgctgtagatttcccggttggagccactgtgtggacataaa 2901  
QY 1924 cgaatgaccccaaggagcagatgcatagtttactcgaactcatatgtagatgagcc 1983  
Db 2902 ctatagagctttagagatgtgtctgtgttagattttagctgtgttcttggctgccaatggcc 2961  
QY 1984 atcgaaaaaacgctgctgttggctcctgaatltccaaagccaaatcggggctgcagat 2043  
Db 2962 gtggggcaagatcagttcaatttgcctcgtacatagcctaagcacaatatacagccac 3021  
QY 2044 ctgttgccttggtagaagaagaacaaatatagacgcgcgctcgaagaaggaaag 2103  
Db 3022 atcatcatgcatltgaaaaaaccccttgaatgtagcgcacacagcagcagcctaag 3081  
QY 2104 ccaagacacatgtagaaggaaatttagatttcgaagaaactccttctctatccatgtgc 2163  
Db 3082 ccgaacacatltggaaggaaatgtacatcttggtaggttagtatactatccacccga 3141  
QY 2164 ccaagatttccatccctcggtgtactatccctcagtaattgagcggagaaacagtagca 2223  
Db 3142 ccgagacatcccaagtgctcagggactgagcctggaggtgaaagagccagacgctgcct 3201  
QY 2224 ttgttggggagcagcgtctgtgggaagaagcattctgttcaacttctgcaggaacttat 2283  
Db 3202 ctgttgggcagcagtggtctgtgggaagacagtggtccagctccctggagcgttctac 3261  
QY 2284 gaccctgtgcaagacagatgctgttgcagtgtgtagatgcaaaagatggaatgtacag 2343  
Db 3262 gacccttggcagaggaaagtgctgtgtgtagtgcgaagaataaagacacatgattgcag 3321  
QY 2344 tggctccgttcccaatagcaatcgttccctcaagagcgtgtgctcttcaactgcaagcat 2403  
Db 3322 tggctccgagacacacctgggcagctgttcccaagagcccatcctgttgcactgcagcat 3381  
QY 2404 gctgagagacatgctgtatgtgtaacaagcgtgtgtgtgcatttagttagtacaagaa 2463  
Db 3382 gctgagaaacttggctttagagacaacagccgggtgtgtgtgtaacaggaagatcgttagg 3441  
QY 2464 gcccaaatgcagaaatatacatcttcttattgaaggttcctcctgtagaatacaacaca 2523  
Db 3442 gccgcaagaaggagcacaatatacatgcttctatcogagcactgcttaataatagact 3501  
QY 2524 caagttagactgaaggagcagccttctgcgcgcgacagaanaaagactagatattgca 2583  
Db 3502 aaagttagagacaaaggaactcgtctctgtgtgccaagaaacaacgcatgtgcatagct 3561  
QY 2584 agggcttcttccaaaaccccaaatattatgtttagatgtaggcacattgaagccctcgat 2643  
Db 3562 cgtgccctttagacagcctcatatttgcctttagatgaaagccagctagactgtgat 3621



Db 2380 aaacgcagagataagtaactgttttccactatgttcttagcccttggaaatttctt 2439  
QY 1324 gtaagtatttcctgagagattattttagcgagagcaggaggaatttaagataga 1383  
Db 2440 attacatttctccctcggggttccacatttggcaagctggagagaccctccacaa 2499  
QY 1384 ttaagacttggcctcacaagcatgtatcatcagagatacttgcgttcttgaataag 1443  
Db 2500 ctccgataacatgtttccgatactcagtcagcagagatgttgagttgttgaccct 2559  
QY 1444 gaaacagcagcagagccttgacaacaattatgccaatagatatagcaaatltaagga 1503  
Db 2560 aaaaaccaccactggagcatgactacacagctcgccaatgatgtcgtcaagttaaggg 2619  
QY 1504 gcaacaggttccagagatgtgcgtcttaacacaatgtcaactcaatcagggaacttcg 1563  
Db 2620 gctataagttccagcgttctgttaattacccaggaataagcaaatcttggacaggaata 2679  
QY 1564 atcattccttaataatgtatggagagatgacattcctgttcttgaattgtctccagta 1623  
Db 2680 attataccttccatctgttgttggcaactaacctgttactcttagcaattgttaccatc 2739  
QY 1624 ctggcgtgacaggaatgtatgtaaacccgcagcaatgtgatttgcacaacaaagataag 1683  
Db 2740 attgcaatacagagagttgttgaatgaataatgtgtctgacacagcactgaaagataag 2799  
QY 1684 caagaaattagcattgtctgaaagatatagcaactgttggaacttgggaatatagactata 1743  
Db 2800 aaaaacttaagaagtgctcgggaagaatcgcctacgtgaagaacaagaactccgcgaacggt 2859  
QY 1744 gtgtcattacaagaaggagaaagccttcagcaaaatgtatgaagaagatgctcagaactcaa 1803  
Db 2860 gtttcttctgactcagcagcagagttggaacatatgtatgtctcagagtttgcaggtacca 2919  
QY 1804 cacaagaatacctcgaaagaagcagcaatatttggaaagctgttatgcatccaagctatgcc 1863  
Db 2920 tacagaactcttcttgaggaagaagcacacattcttgaattacatttcccttaaccaggca 2979  
QY 1864 ttatatttcttctctctgacagcagggttccagatttggagccttaattcaagcctgga 1923  
Db 2980 atgactgtatttcttccatgcctgagatgttcccggttctggagcctcctgtgtgacacataa 3039  
QY 1924 cgaatgcaccagagagcagtgatcatagtttcttactgcaattgatalgtgagctatggcc 1983  
Db 3040 ctactgagcttggagagatgtctgttagtatttccagctgtgtcttggcagctatggcc 3099  
QY 1984 atcggaaaaacgctcgttctgtgctcctgaataatccaaagccaaatcgaggcgctgcgcat 2043  
Db 3100 gtggggaagtcagttcatctgtcctcgtacatgtccaaagccaaatatacagcagccac 3159  
QY 2044 ctgttctcctgttggaaagaagaacaatatatagacagcgcgagtcgaaggaaggaataag 2103  
Db 3160 atcatcatgacatctgaaaaaaacccttgaattgacacatcaacacggaagccctaatg 3219  
QY 2104 ccagacacatgtgaaaggaatttagagtttctgagaagctcttcttctatccatgtcgc 2163  
Db 3220 ccgaacacatgtgaaaggaataatgtaacatttggtagaagtgttagtaactataccacgga 3279  
QY 2164 ccagatgttctatccctcgtgtgcttactccctcagttatgtgacgaaggaagacagtagca 2223  
Db 3280 ccgagacatcccaatgtctcagagactgagcctggaagtgaagaagggccagacgtgct 3339  
QY 2224 ttgttgaggagcagcagctgtgagggaagaagcactctgttcaactctgcaagaacttat 2283  
Db 3340 ctgttgaggcagcagtgctgtgagggaagcagctgttccagctccctggaagcgttccac 3399  
QY 2284 gaccocgtgcaagacaaagtctgtttagatgtgtgtagtgcacaagaatltgaatgtacag 2343  
Db 3400 gaccoccttgcaaggaaagtgtctgttagtgcacaagaatataagacgacatgaattgtccag 3459  
QY 2344 tggctcgttcccaaatgaacatcgttctcctaagaagcgtgtgtcttcaactgcagcaat 2403

Db 3460 tggctccagacacacactgggcacgtgtgtcccaaggagcccatcctgttactgcagcatt 3519  
QY 2404 gctgagaacatcgcctatgtgtgacaacagcgtgtgtgtccattagatgagatcaaga 2463  
Db 3520 gctgagaacatcgtcctatgtgagacaacagcgtgtgtgtccattagatgagatgtgag 3579  
QY 2464 gccgcaatgcaagaataatcatcttcttatttgaaggtctcctctgagaataatacaca 2523  
Db 3580 gcagcaagaagagccacacatacatgtccttcatcagatcactgtcctaataataatagcact 3639  
QY 2524 caagttgagactgaaagagacagacttctgcggccagagaacaaagactagctatgtca 2583  
Db 3640 aaagtgaagagacaagaagaaactcgtcctgtgtgcagagaacaaagcattgtccatagct 3699  
QY 2584 aggccttctcccaaaaaccacaaatttattgttgaatgagccacttccagcctcagat 2643  
Db 3700 cgtgccttgttgaacagcctcatatttcttcttgaatgagcagcgtcagctcgtgat 3759  
QY 2644 aatgacagtgaagaaggtgtgtgacgcatgccccttgataaagccagagcaggaagagacatgc 2703  
Db 3760 acagaaagtgaagaagtgtgtccaaagagcccttgacaaagccagagagagccgcacttgc 3819  
QY 2704 ctatgttctactcaagcgtcctctgcgaattcagaagcagagattgagtggtctgcac 2763  
Db 3820 atgtgtatgtctacacgcgtcgtccacacacacacacacacacacacacacacacacac 3879  
QY 2764 aatgaaagataaagagacaagagaaactcatcaagaagcctccttgagaatctcgagacataat 2823  
Db 3880 aatgagcagatcaagagagacatgtgcacgcatcagcagactgtgtgcagagaagcattcat 3939  
QY 2824 tttaagttagtgaatgcacagtcag 2848  
Db 3940 ttatcaatgtgcaggtgtccaggtcgt 3964

RESULT 11  
AAV6534  
ID AAV6534 standard; cDNA; 4264 BP.  
XX  
AC AAV6534;  
XX  
DT 08-JAN-1999 (first entry)  
XX  
DE Mutated human P-glycoprotein gene designated mdr-delta-F336.  
XX  
KW mdr1-delta-F336; cyclosporin-resistant deletion mutant; human;  
KW P-glycoprotein; cyclosporin; drug efflux;  
KW modulator-resistant multidrug resistance; cancer therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN US5830697-A.  
XX  
PD 03-NOV-1998.  
XX  
PF 21-JAN-1997; 9705-0784649.  
XX  
PR 21-JAN-1997; 9705-0784649.  
XX  
PA (STRD ) UNIV LEIAND STANFORD JUNIOR.  
XX  
PI Chen G, Slik BI;  
XX  
DR WPI; 1998-609244/51.  
XX  
PT DNA encoding mutant P-glycoprotein polypeptide - useful for  
PT producing recombinant polypeptide  
XX  
PS  
XX  
XX Disclosure; Columns 27-32; 17pp; English.  
CC The present sequence represents mdr1-delta-F336, a gene encoding a  
CC cyclosporin-resistant deletion mutant of human P-glycoprotein lacking  
CC Phe336. The recombinant protein can be used to produce antibodies. It

QY	1046	-----agtcagcttcattgacaaggtctgaggaatccaaccatct	1086
Db	2140	actcgtagagtgctgcgttgatcaacaagcccaagaagaagctctagtaaccaagaagc	2199
QY	1087	aagaagataagctctccttgaaagctctctctattuaaaattttaaagtttaacaagcttca	1146
Db	2200	ctgagatgaagataaacctccaagtttcctcttggagattatgaagctaaatttaactgaa	2259
QY	1147	tgcgctcttgctggtctgcgggagacattgcgtctgctctaaatggaactgctatccagta	1206
Db	2260	tgcgctattttgctgtgtgtgtatttttggcattataaagggggcgtgcgaaccagca	2319
QY	1207	tttccatcactcttgcacaaaattaaaccaatgcttggaaataatgataaaac---aca	1263
Db	2320	tttgcataataattttcaagaattataagggtttcttacaagaattgatactccctgaaca	2379
QY	1264	ttaaagatgatgcaagaattatttccatgatatctgcattttgggtggtatttgctt	1323
Db	2380	aaacgaacgaattagtaactctgttttccaactatgtttctcagcccttggaaattttcttt	2439
QY	1324	gtcagttatttctcagccagggattattttaocgcagagcgaggaaattttaacgaatga	1383
Db	2440	attacatttttccctccagggttccacatttggaaagcctggagagatccccaacagcg	2499
QY	1384	ttaaagacacttgcctccaagaacgatgtatatacagagatatgctcgtgttgatgaaag	1443
Db	2500	ctccgataacatggtttctcgcatacctacatgctcaagcagatgtagtgtgttatgacct	2559
QY	1444	gaaacacgcacaaagagcttgcacaacatatgagccatgatatgacagcaaatatgaaga	1503
Db	2560	aaaaacccactctggagcattgactaccacggctcgcacaaatgactgtcctaagttaaag	2619
QY	1504	gcaacaggttccagagatttgcgtccttaacaacaaatgcgaactaaatggaacttcagt	1563
Db	2620	gctataggttccagcgtcctgcgttaattaccagaaatatgcacaaactctgggacagaa	2679
QY	1564	atcatttccttatatatggaatggagagatgacattccgatctcagatgtgtccagta	1623
Db	2680	attatattcctcatcatgattgttggcaacataacactgttactcttgaactgttaccact	2739
QY	1624	cttgcgctgaaaggaatagattgaaacccgcgcacatgactcggatttgcacaagaataag	1683
Db	2740	attgcataagcgaagagattgtttgaaatgaaaatgttctctgcgcacagcacatgaaataag	2799
QY	1684	caagaacttaagcagctcttggaagaatagcaacgtgaagctcttggaaataagtgctcta	1743
Db	2800	aaagaacttagaagtgctctgggaaagatccgcctaccggaacatagaanaacttccgaacgt	2859
QY	1744	gtgtcatataaagaaggaaaaagccttcagacaaatglatgaagagatgcttcaagctcaa	1803
Db	2860	gttctcttgactcagggcagaagatttgaacatatgtatgtccacagatttgaggtaacca	2919
QY	1804	cacagaaatacctcgaaagaagacagatatattggaagctgttatagcattccagcatgcc	1863
Db	2920	tacagaaactcttggaggaagaacacacatcttgggaattacatacttccctccacaggca	2979
QY	1864	tttatataatttgctctatgacagaggttgcagtttggagtcctatttaattcaagcttga	1923
Db	2980	atgagatatttctccatgcctgagatgttcccggtttgggagccctactgtgtgcacaataa	3039
QY	1924	cgaatgaccccaagggcagcttcaatagtttcttactgcacattgcatatgagactatgccc	1983
Db	3040	ctcaagacttggagatgtctgttgaatatcttcaagcctgtgtccttgggtgcacgtgccc	3099
QY	1984	atcggaanaacgctcgttttggctcctgtaatactccaagcnaactcggggctgycgat	2043
Db	3100	gtggggcgaagtcagttcaattgtctccctgactatgccaagaacaaaatalcacagcagccac	3159
QY	2044	ctgtttcgtcgtgttggaagaagaacaacatatgacagccgcagtcacgaagaaggaaag	2103
Db	3160	atcatalcatgataltgaanaaaaccccttgatgtgacagctacacgaacgaaagcctaactg	3219
QY	2104	ccagacacatltgaagggaatttagaatttccggaagctctcttctcatcaccatgctgc	2163

DB 3220 ccgaacacatggaagaaatgcatcatttggtgaagtgtatcctaactatccaccga 3279  
OY 2164 ccagatgtttatcatctccgtggtctatccctcagatattgaagcgagaaagcaataga 2223  
DB 3280 ccgagatcccaatgcttcaaggactgagccgtgaagtgaagaagggccagacgtcgtc 3339  
OY 2224 ttgtgggagagcggctgtgggaaagcaactctgttcacactctcgaagacttat 2283  
DB 3340 ctggtgggagcagcagctggtgtgggaaagcagatggtccacgtcctcgtgagcgttccac 3399  
OY 2284 gaccccggtgaagcaagatgctgtttgtagtgtgtagtgcacaaagaattgaattagac 2343  
DB 3400 gacccttgtagcaggaagatgctgtctgtagcagaagaataaagcagactgatttcag 3459  
OY 2344 tggcttcgttcccaatgcaatcgtctcctcaagagcctgtgtcttccacgcagcat 2403  
DB 3460 tggcttcgagcaccactgtggtcgtgtccagagcccatcgttltgactgcagcat 3519  
OY 2404 gctggaacatcgcctatgtgtacaaacagcgtgtgtgtcctatgattgataaaga 2463  
DB 3520 gcttggaacatgtcctatgtggaacaaacagcgtgtgtgtcctcagaagaagattgtgag 3579  
OY 2464 gccgcaatgtagcaaatatccattcttattgaagtgtccctcgtgagaataacacaca 2523  
DB 3580 gtagcaaaagagccaacatacatcctcatcagtcactcgtcaataataatagcaact 3639  
OY 2524 caagtgtgactgaagagagcagactttctgtgcgcagcaaaacaaagatactattga 2583  
DB 3640 aagatgtaggagacaagaactcagctcgtgtgtgcagaaacaaacagcattgcatagct 3699  
OY 2584 aaggctctctccaaacaaacaaatttatctgtgtgtagtaggccaactcagccctcag 2643  
DB 3700 cgtgcctcgttagacagcctcatattttgtcttggatgaagccacgtcagcttgcgt 3759  
OY 2644 aatgacagtgaagagtgtgttcagcatgcccctgtgtaaaagccagagcgggaagcaatgc 2703  
DB 3760 acagaaagtgaaaaggtgtgtccagaagccctggaacaaagccagaaagccgcacctgc 3819  
OY 2704 ctatggtgtactacacaggtctctcgaattcgaacgcagcagatttgatgttctgcac 2763  
DB 3820 attgtatgtctcaccgcctgtccaccatccagaaatgcagacttattatgtgttccag 3879  
OY 2764 aatgtaaaagataaagaaacaaagaaactcaagaagcctcctcagaagaatcgaacatata 2823  
DB 3880 aatgtagcagatgaagagatgtagcagcgtcagcgtcgtgtagcagaaagcagcatctat 3939  
OY 2824 tttaagttagtgaatgcacagtcag 2848  
DB 3940 ttccaatgtcagtgctcagagctg 3964

## RESULT 12

AAZ24042 ID AAZ24042 standard; CDNA; 8630 BP.

AC AAZ24042;

DT 04-FEB-2000 (first entry)

DE Retroviral vector SPbeta9ImSAI mdr-1 CDNA.

KW Retroviral vector; gag gene; gene therapy; chemotherapeutic agent; hematopoietic stem cell transformation; mdr-1; ss.

XX Retrovirus.

OS EP955374-A2.

XX 10-NOV-1999.

PF 07-MAY-1999; 99BP-0250151.

XX

PR 08-MAY-1998; 98DE-1022115.  
XX  
PA (PETT-) PETTE INST HEINRICH.  
XX  
PI Osterlag W, Baum C, Hildinger M;  
XX  
WP1: 2000-001087/01.  
DR  
XX  
PT New retroviral vector containing minimal or no gag gene sequence, for  
XX use e.g. in gene therapy or cloning  
XX  
PS Disclosure: Page 19-21; 35pp; German.  
CC  
CC This invention describes a novel retroviral vector (RV) containing a gag  
CC gene fragment having fewer than 400 bp. RV are used: (1) in gene therapy;  
CC (11) for cloning genes; (111) for (over) expression of proteins or RNAs  
CC and (1V) for transfection of chemotherapeutic agents). Because of their reduced  
CC content of viral genes, RV are very safe (no expression of toxic or  
CC immunogenic proteins, no recombination with other viruses), have  
CC increased cloning capacity and express non-viral sequences at a high  
CC level. This sequence encodes a retroviral SPbeta9ImSAI mdr-1 protein  
CC which is described in the method of the invention.  
XX  
SQ Sequence 8630 BP; 2234 A; 2011 C; 2189 G; 2196 T; 0 other;

Query Match 36.0%; Score 1028.6; DB 21; Length 8630;  
Best Local Similarity 62.2%; Pred. No. 9.4e-263;  
Matches 1694; Conservative 0; Mismatches 983; Indels 42; Gaps 3;

OY 166 attatgtaaaagccagatagataactttccacagctgagatataaactgatacc 225  
DB 2321 ataattgtaataaagccagatattgacagctatttgaagagtgaggcacaacacagataat 2380  
OY 226 atagaagaactgtgaaatttaaaatgttcttcaattatccatcaagacatctc 285  
DB 2381 attaagggaatttgaaattcagaatgttccactcagttaccatccatccgaaagaagt 2440  
OY 286 aagatctgaaagtcgaatcgaatcgaattgaatgctggaagagcagctgtgtcgt 345  
DB 2441 aagatctgaaagtcgaatcgaatcgaattgaatgctggaagagcagctgtgtcgt 2500  
OY 346 ctcaatgtagcagtggaagagtagcagtagtccagcttccgagaggttatalatgacccgat 405  
DB 2501 aacagtggtgtggaagagcacaacagtcagcgtcagcagaggtctatagaccaca 2560  
OY 406 gatggtctatcatggtgtagaataatcatcagcagcttcaatgtcgcgcatatga 465  
DB 2561 gagggagatgtcaggtttagtagaagatatagaccataaatgaagtttctacgg 2620  
OY 466 gaccatattgaggtgttgaagcagcgtgttcttcggagccacatcagtaaac 525  
DB 2621 gaatacattgtgtgtgtagaagaaactcgtatgttgcaccacagatgctgaaac 2680  
OY 526 atcaagtagcagagatgtagtgcatagagaagatgtagagagcagcaaggaagca 585  
DB 2681 attcgtatgtagcagagatgtagtgcatagagaagatgtagagagcagcaaggaagca 2740  
OY 586 aatgctatgatttataatcagagtttccataaatttaacttgtgtgggaaaaa 645  
DB 2741 aatgctatgatttataatcagagtttccataaatttaacttgtgtgggaaagaa 2800  
OY 646 ggaagtcgaatgtaggagagcagaaagagagatcagcaatcgtcgtcgtcgtttagca 705  
DB 2801 ggggcccagtgagtggtgtagcagaagcagagatcgcacatgcagcgtgctgttcgc 2860  
OY 706 aaccccaagatctgattatgtagagcagctcgtccctgtagtccagaaagcaatga 765  
DB 2861 aaccccaagatctcctcgtgtagagcagcagcagcagcgttgcagcagaaagcaagca 2920  
OY 766 gctgttcaagctgcaactgtagaagcagcaaggtcgtgagctacatcgtgtgtagcaac 825



D 2921 gtgttcaggtgtgctctgtatagaagccagaagaagctcgagaccacatgtgatagtcat 2980  
QY 826 cgacttctactatccgaagtgcagattgtgtgacccctaagatgtaagtcgtcgcg 885  
D 2981 cgttctgctacagcttcgttaatgcgcgcacatcgctgtgttcgatagtgagtcattg 3040  
QY 886 gagaagaagcagcagctgtaactatgcaaacgagctctatattcatctatgtagt 945  
D 3041 gagaaaggaatcatcagatgtaactatgaaagaagaagcatttacttcaactgttcaaa 3100  
QY 946 tcaacgagatataaagaagctgataacagatgtagtcaatgacatatctta----- 997  
D 3101 atgcagacagcaggaatgtaagtattgataaataatgcagctgatacccaaaagtga 3160  
QY 998 -----ctgaaagaagaagaccactcctcctgtgacactgtgtaagaacatca---- 1045  
D 3161 attgattcgttgaaatgtcttcaaatgtattcaagatccagcttcaataagaanaagatca 3220  
QY 1046 -----agtcagacttcatgacaagcgtcgaggaatcccccacatct 1086  
D 3221 actcgttagagtgctcggtgagtcacagaagccagaagaagaagcttagtaccagaagct 3280  
QY 1087 aagaagataagctctccgtgaagctctctatataaanaatttaaglttaacaagcctga 1146  
D 3281 ctgagtgaaagtataccctccagcttctccttctgagagattatagaagcttaattcaatga 3340  
QY 1147 tggccttctgtgtctcgtgggacattggtctgttcttaataaggaactttatccagta 1206  
D 3341 tggccttattgtctgtgtgtgtatttctgcatataaataatgagggcctgtaacacaga 3400  
QY 1207 ttcttcacatcttctgcaaaaataataacacatgcttgaataatgataaataaacc---aca 1263  
D 3401 ttgtcaataataatttcaagaagttataggggttttaacaagaattgatactcctaataaa 3460  
QY 1264 ttaagcagatgtagcagaataattatccatgatactcgtcaatttgggtgtatttggctt 1323  
D 3461 aaacagcagaatagtaactgttcttcaactatgttctcagcccttggaaattttcttctt 3520  
QY 1324 gtcagttattcatgacaggaattatttcaaggcagagcgagggaattttaaagatgaga 1383  
D 3521 attcaacttctcctccaagtttccacatttgcgaagcgtcgagagctccctccacaagcg 3580  
QY 1384 ttaagacacttgccttcaagccatgcttataacaggaatctgacctgttatactgaaga 1443  
D 3581 ctcgataatactgttctcgcagatcagtcacagacagagatgtagtgggtttagtgacct 3640  
QY 1444 gaaacagcagagaggtctgcaacaatatagccatagatatagcacaaattcaaga 1503  
D 3641 aaaaacacacactgagcattgataccaggtcgcgaatgactgctcgaattaaagg 3700  
QY 1504 gcaacaggttccagatgtgcttcttaacacaaatgcaactaaccttggacttcaagt 1563  
D 3701 gctataagttccaggtctgttlaattlaccagaatagcaaatcttggscagaggaata 3760  
QY 1564 atcatcttcttataatgatgagatgacatccctgattctgtgatacttccagta 1623  
D 3761 attatactctcatcatatgattgtgcaactaacacgttactcctttagcaattgtaccac 3820  
QY 1624 ctgtgcgtgacaggaatgattgaacccgagcaatgactggaatttgcacaacaagaataag 1683  
D 3821 attgcatagcagaggttctgtaaatgtaaatgttctggaacagcactggaagaagtaag 3880  
QY 1684 caagaacttaagcagctggaagaatagcaactgaaagcttggagagataagtaactata 1743  
D 3881 aagaacttaagaggtgtgtggaagatcgtactgaaagcaatagaanaacttccgaacgct 3940  
QY 1744 gttcatataacaggaanaagcctcgacaatgtatgaaagagatgcttcagactca 1803  
D 3941 gttcttctgactcagagagagaagtttgaacatatgtatgcccagaggttgcaggtacca 4000  
QY 1804 cacagaataactcgaagaagaacagagattatggaagctgtatgcatctcagcatalgcc 1863  
D 4001 taagaactcttggaggaagaacacacatcttggaaattatcttcttccaccagga 4060

QY 1864 ttataatatctctatgcaagcagaggttctgattggagacttatttaacagtga 1923  
D 4061 atgataatttctccatgctggaatggttccggttctggacctacttgtgccaataaa 4120  
QY 1924 cgaatgaccccaagaggcctgttcatagtttcttactgcaattgcatatggaactatggcc 1983  
D 4121 cctatgagcttggagatgttctgttagttatttccagctgttcttgggttgcattggcc 4180  
QY 1984 atcggaaaacgctcgttctgtcctgtaatatcccaaaagccaatctcggtcgctgcgat 2043  
D 4181 gctgggcaagtcattcatcttctcctgactatgccaagaacaaatatacagagccac 4240  
QY 2044 ctgttgccttcttggaaaagaacaaatatacagacagcagtcgaagaagaagaag 2103  
D 4241 atcatcatgatactgtaaaaaccccttggattgacagctacacaggaaggtcctaattg 4300  
QY 2104 ccagacacatgtagaagggaatttagagtttcgagaagctcttcttctcatcatgtcgc 2163  
D 4301 ccgaacacatggaaggaatgtaacatttggtaggtgtatcaactatccaccgga 4360  
QY 2164 ccagatgttttcatcctcgtgtgttatacctccagatgtagcaggaagaagatagca 2223  
D 4361 ccgagacatccagtgcttcaaggagctgagcctggaagtgaagaagggccagacgctgtgc 4420  
QY 2224 ttgttgggagagcagcggtgtggaaaagcacttctgtcacttctgcaagacttatt 2283  
D 4421 ctgtgtggagcaggtgtgtgtgggaagacaagtgttccagctccctcgtgaagcgttctac 4480  
QY 2284 gaccccgctcagaaggaagaatgctgtgttgaatgtgtgtagatgcaaaaagatltgatagca 2343  
D 4481 gaccccttgagaggaaggtgtgtgtgtgaatggaagaagaataagagctgatagttcac 4540  
QY 2344 tggctcgttctccaaatagaatcgttctcctaagaagcctgtgtcttcaactgcgcagcat 2403  
D 4541 tggctcggagcacacctgtgcatcgtgtccaggaagcccatcctgttctgactgcagcat 4600  
QY 2404 gctggaacatcgccatgtgtgacaacagcggtgtgtgttcattatgatactgaagaaga 2463  
D 4601 gctggaacatgtgctcatgtagagaacaacagcggtgtgtgttcaagaagaagatcgtgaag 4660  
QY 2464 gccgcaaatgcaacaatatccattcttattgaaggtctccctgagaagaataacaaca 2523  
D 4661 gcaagcaaggaagccaacatacatatgcttcatcgtgactgactaataatagcattac 4720  
QY 2524 caagtgtgactgaagagacagcttctcgtgcgcagaaacaagaactagctatgca 2583  
D 4721 aaagttagagacaagaagaaactcagctcctgtgtgtgcagaagaacaacacgcatltgcatagct 4780  
QY 2584 aggcctctctccaaaaccaaataatttattgttggatggagcgaccttcagccctcgat 2643  
D 4781 cgtgcctctgttagacagctctcatatttgccttcttggatgaagccgcgtcagcttggat 4840  
QY 2644 aatgacagtgaagaagtgtgttcaagctgcccctgtgaataagccaggaaggaagacatgc 2703  
D 4841 acagaagaagtgaagaaggtgttccaagaagccctgtgacaagaagccagaagaagccgacactgc 4900  
QY 2704 ctatgtgtcactcaacaggtctctcgaattcagaacgcgatttgaatggttgcgcac 2763  
D 4901 attgtattgtcacaacgctgttccacacatccagaatgacgacttaagtgtgttcag 4960  
QY 2764 aatggaagaatagaagaagaagaactcatcaagaagctcctcgaatctggaatctggaacatatat 2823  
D 4961 aatggaagaatgcaagaagagatgcaagcattcagcagctgtgtgcacagaagaagcattat 5020  
QY 2824 tttaagttagttgaatgcacagtcag 2848  
D 5021 ttctcaatgtcaggtgtccaggtctg 5045

RESULT 13  
AA294742  
ID AA294742 standard; cDNA; 3924 BP.



XX AA294742;  
AC 01-AUG-2000 (first entry)  
XX  
XX  
XX  
DE Human ATP binding cassette ABCB4 (MDR3) cDNA.  
XX  
XX ABCB4; ATP binding cassette; human; cholesterol; lipid disorder;  
XX atherosclerosis; lipid disorder; dyslipidemia; psoriasis;  
XX lupus erythematosus; diagnosis; gene therapy; MDR3;  
XX multidrug resistance; chromosome 7q21; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200018912-A2.  
XX  
XX 06-APR-2000.  
XX  
XX 21-SEP-1999; 99WO-EP06991.  
XX  
XX 25-SEP-1998; 98US-0101706.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Schmitz G, Klucken J;  
XX  
XX WPI; 2000-293151/25.  
XX  
XX Adenosine triphosphate binding proteins useful for identifying agents  
XX PT for treating atherosclerosis and other inflammatory disorders -  
XX  
XX Claim 9; Page 117-119; 154pp; English.  
XX  
XX  
XX The present sequence is that of human ATP binding cassette  
XX subfamily B protein ABCB4 cDNA. The cDNA was identified using a  
XX differential display method in which monocytes from peripheral  
XX blood were subjected to macrophage differentiation and cholesterol  
XX loading with acetylated low density lipoproteins and subsequent  
XX deloading with high density lipoprotein (HDL3) to identify  
XX cholesterol sensitive genes. The gene maps to chromosome 7q21  
XX and is also termed MDR3 (multidrug resistance). The invention  
XX provides cholesterol-sensitive ABC genes (see AA294734-63). These  
XX genes, and polypeptides encoded by them, can be used for diagnostic  
XX and therapeutic applications, and for biochemical or cell-based  
XX assays to screen for pharmacologically active modulator compounds  
XX useful for the treatment of lipid disorders, atherosclerosis or  
XX other inflammatory diseases such as psoriasis and lupus  
XX erythematosus.  
XX  
XX Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other:  
XX

Db 1380 gaagagcaacaattacattgatgagcagatagatagaactttaatgtaactatccgaag 1439  
QY 466 gaccatattgagatggttagtcaagaagccgtgtttgtcggaaccacatcagtaacat 525  
Db 1440 gaatcatgtgtgtgtagtcaagagccggtgtctgttccaccacaattgtctgaaat 1499  
QY 526 atcaagtatgagagatgtagtgaactgtagaagaatgtagaggaagcaagggagaaga 585  
Db 1500 attgtatgcccgtggaatgttaacatgtagagatlaaagaagctgcaagaagggcc 1559  
QY 586 aatgcgtatgatttatacatgaggttccctaataatttaatacatggttagggaaaaa 645  
Db 1560 aacgctatgagtttatcatgaataattaccagaataattgacacccctggttggagaga 1619  
QY 646 ggaactcaaatgagtggaaggcagaanaacagagatcgaattgctcgtgcttaattga 705  
Db 1620 ggggcccagctgagtggtgagcagaagcagaagatcgcaattgacggtccgtgtcgc 1679  
QY 706 aaccccaagatctgattttagatgaggtacgtctgcacctgtagatccagaagaagtc 765  
Db 1680 aaccccaagatctctctgctgtagtaggcccagctgcagcatltagacacagaagtgagct 1739  
QY 766 gctgtcaagctgcaactggaagagcgagcaaaagctcgactacaactcgtgtagcacac 825  
Db 1740 gaggtaacagcagctcgtgataagggccagaagaagccggacacacatgtgatagcacac 1799  
QY 826 cgaacttctactatcgaagtgcagatltgattgtgaaccctaagaatgtagatgtgcg 885  
Db 1800 cgaactgtactagctcgaagaatgtagatgtcatcgtgtgtttagagtagagtaattgtg 1859  
QY 886 gagaagaagacatgctgtaacctatgccaanaacgaaggtcatatattcatcttgat- 944  
Db 1860 gaggcaaggaagcacaagcagacatgagtagaagaagggtgtgtaattcaactgttcaac 1919  
QY 945 -----gtcacagatattaaaaagctgtagacaagatg 978  
Db 1920 atgcagacatcaggaagccagatccatcctcgaagaatttgaacaaatgltgaaagct 1979  
QY 979 gagtcaatgacatatctactgaaagaagaaccaactcaactcctctgcactctgtgaag 1038  
Db 1980 gccactgaatggtcccaaatggtgtagaatctgcctattttagcacttactacagaaga 2039  
QY 1039 agctcaagtcagactcattgacaagc-----ttaggaatccaccacaactcaagaag 1092  
Db 2040 aaccttaaaatttacaagaatgtcagaagaagccttgaatgtagaaccagatggaactgaa 2099  
QY 1093 ataagcttcctgaagctcctcctaataaaatttaaaagtttaaacagaagcctgaatggcct 1152  
Db 2100 gcaaatgtgccacagatgctccttctgaaggtcctgaacatgaaataaacagatggccc 2159  
QY 1153 ttgtgttctggyggaattggtcttctgtaaaatggaactgttcatcagatatttcc 1212  
Db 2160 tactttgtcgyggaacagatgctgcatlgtccaatgtggggtcctcagcggcatttca 2219  
QY 1213 atcattctgaaaaaattataaacatggttggaaataatgataaacacaacttaagaat 1272  
Db 2220 gtcatattctgaagatataagcatttctggaacaggcagatgtagcagtgagaagcagag 2279  
QY 1273 gatcagaatatttcatcatatctgcatactgttgggtgtgtattgttgcattat 1332  
Db 2280 aagtgcaacatatctctttagtttcttattcttcttcttcttcttcttcttcttcttctt 2339  
QY 1333 ttcatgcaaggatattttagcgcagagcagggaaatttaacgatgagatgaagacac 1392  
Db 2340 ttccctcaggggttcaagcttgggaagcctgycagatccctcacagaagaagactcggtlca 2399  
QY 1393 ttggtctcaagacagatgtatatcagagatgtgcgtgttggtagaaagaaagaaagc 1452  
Db 2400 atggttttaagaatgttaagaacagatgactggtgttggtagacaaataaacagct 2459  
QY 1453 acagagagcttgacaacatatagccatagatatagacaataatcaaggaagcagat 1512

Db 2460 actggtgacttcttacaagacttgcacagatgctgcaccaagtcaccaagagcacagga 2519  
Oy 1513 tccaggtatgctgttctaacaacaatgcactaacaatgagacttccgtatcttcc 1572  
Db 2520 accgagttggtctttaatgacacagaatataagcttaacacttggaactgtattctatca 2579  
Oy 1573 ttataatagtgaagatgacattcctgattctcgatgttcgagttgcccagttacttcgtg 1632  
Db 2560 ttatctcgttctggtcagattcaacctattgtctattagcagttgttccaatatttctgtg 2639  
Oy 1633 acagaatgattgaaacccgcgcgaatgactgtatttgcacaacaagaataaagcaaat 1692  
Db 2640 tcaagaaattgttgaatgaatgttgtgtggaatgccaaaagagataaaaaagaactg 2699  
Oy 1693 aagatctgtgaagaatgacactgtgaacttcttgagaataaactgactatagtcattg 1752  
Db 2700 gaagctgtggaagaatgtgcacacagagcaatagaataatagaacagttgttctgtg 2759  
Oy 1753 acaagggaaaagaccttcgagaatgtatgaagagatgcttcagaccacaacagaat 1812  
Db 2760 acccaggaagaagaatttgatcaaatgtatgttgaataatgtatgacacttaacaggaat 2819  
Oy 1813 acctcgaagaagacagactatttgaagctgttatgtacatcagcacttcttataat 1872  
Db 2820 tctgtgcagaagacacacatctatggaattactttagtactcacaagcattatgtat 2879  
Oy 1873 ttgcctatgacagcaggttctcgatttgagcctattcaattcaagctggagcaatgacc 1932  
Db 2880 ttctctatgctcgtgttcttcgatttctgcatattcattcattgtgaatgacatttgcgc 2939  
Oy 1933 ccaagggagcatgtcatgttcttactgcaattgcatattgagctatgcacatcgaa 1992  
Db 2940 ttccagagttgtatctctgtgttcttcgcaattgtatttgtgcagctgtgcttaagacat 2999  
Oy 1993 aagctcgttctgtgctcctgaatattccaagccaatcgggggctgtgcacatgttgc 2052  
Db 3000 gccagttcatctgcacagactatgctaagctgtctgcagccactattctatg 3059  
Oy 2053 ttgttggaagaagaacccaatatagacagccgcagtcagaagaaggaaaaagccagaca 2112  
Db 3060 ctgtttgaaagaacaacctctgatttgacagctacagtgaaagggtctgaagcctatataa 3119  
Oy 2113 ttgtgaagggaatttgaagttcttcgaagaagctctcttctcttaccatgtcccgcaatgt 2172  
Db 3120 ttggaaggaataataacatttaatgaagctgtgttcaactatccaccggaacaagctg 3179  
Oy 2173 ttcatctcctgtgtgcttaccctcagttatlgagcgaaggaagacagtagcatlttg 2232  
Db 3180 ccagtgctcaggggtcgtgagccttgaggttgaaagaagccagacactagccctgtgtg 3239  
Oy 2233 agcagcgtctgtgggaaagcaactctgttcaactctgcagaagacttatagcccgtg 2292  
Db 3240 agcagtgctgttggaagagcagctgtgtccagctccctgagcgttctacagaccctgtg 3299  
Oy 2293 caagagcaagagctgttgaaggtgttgatgcaaaagaattgaatgacagtggtcct 2352  
Db 3300 gcggggaagagctctctcgatggttcaagaagaagaacaactcaatcgtccagtggtcaca 3359  
Oy 2353 tcccaaatagaacatcgttctctcaagagcctgtgtcttcaactgcagatgtctgagaa 2412  
Db 3360 gctcaactcgaatcgttctcgaagagcctatcttattgttagcagatgtcgggaat 3419  
Oy 2413 atgcgtatgttgaaacaagcgtgtgtgtccattagatgagatacaagaagccgaat 2472  
Db 3420 attgcttatgagacaacagcgggtgttaltacagagatgaattgtgagtgacaccaa 3479  
Oy 2473 gcaagaatatcattcttatttgaaggtccctgagaagatacaacaacaagaatttg 2532  
Db 3480 gctgcacaacatacattccttctatctgagagcttaaccacaataatgaaacagagttgg 3539  
Oy 2533 ctgaagaagagcagcttctctgtgcggcagaagaacaagaactgattgtcgaaggctct 2592  
Db 3540 gataagggagctcagctctcagggagtgcaaaaacagagattgtctattgcccggagccct 3599

Oy 2593 ctccaaaaaaccaatttattgttggatgagggcacttcagccctcgaataatgacagt 2652  
Db 3600 atcagacaacctccaatccctcctgttgatgaagctaacacagctctctgatactgaagt 3659  
Oy 2653 gagaaggtgttctcaacatgcctctgataaagcagagcgggaagagacatgctatgtgc 2712  
Db 3660 gaaaaggtgttccaagaagcctctgacaagaagcagagaagggccgacccgacttggatt 3719  
Oy 2713 actcaggtctctgcgaattcagaacgcagatttgatgtgttctcgcacaatggaag 2772  
Db 3720 gctcacggcctgtccacccatccagatgcagacttaatagttgttccagatggaga 3779  
Oy 2773 ataagaagaacaggaactatcaagaagctccctgagaatccggaacataatttaagta 2832  
Db 3780 gtcaaggaagcatgtgcagcatcagcatgctgtgcacagaagaagcatctatttcaatg 3839  
Oy 2833 gtgaatgcacagtcagt 2850  
Db 3840 gtcagtgctcaggtcgg 3857

RESULT 14  
AAZ88974  
ID AAZ88974 standard: DNA: 3924 BP.  
XX  
AC AAZ88974:  
XX  
DT 30-MAY-2000 (first entry)  
XX  
DE Human MDR-3 DNA.  
XX  
KM Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;  
KM stomach; colon; ulcerative colitis; pouchitis; large intestine; human;  
KM inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;  
KM ileum; anti-ulcer; anti-inflammatory; cyostatic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200007577-A2.  
XX  
PD 17-FEB-2000.  
XX  
PF 06-AUG-1999; 99WO-DE02426.  
XX  
PR 06-AUG-1998; 98DE-1035526.  
PR 15-DEC-1998; 98DE-1057750.  
XX  
PA (STRE/) STREMMEL W.  
XX  
PI Stremmel W;  
XX  
DR WPI; 2000-195439/17.  
XX  
PT Using phosphatidylcholine having mucous membrane protective activity to  
PT prevent or treat large intestinal diseases, such as ulcerative colitis  
PI or pouchitis -  
XX  
PS Disclosure; Page 17-22; 22pp; German.  
XX  
CC This invention describes a novel medicament containing a therapeutically  
CC effective amount of phosphatidylcholine to treat diseases, where  
CC phosphatidylcholine has an advantageous mucous membrane protective  
CC activity in the large intestine. The invention describes (1) a method  
CC to regulate presence or absence of MDR(multidrug resistance)-3  
CC transcription in a sample from patients with suspicion of stomach and  
CC colon diseases, where treatment with phosphatidylcholine is indicated;  
CC (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large  
CC intestinal inflammation, Crohn's disease, diverse colitis, infectious  
CC enteritis/colitis, inflammation through X-ray treatment, antibiotics,  
CC chemotherapeutics, drugs or chemicals or large intestinal carcinoma,  
CC contains MDR3-specific primers to detect mutations in the MDR3 gene  
CC family and their interactions partners (in particular transcription

CC factors), and (3) a method to manufacture a medicament for local gene  
CC therapy in the ileum with the ideal MDR3 analogus genes for patients  
CC with missing or decreased expression of MDR3-analogous proteins. The  
CC products of the invention have anti-ulcer, anti-inflammatory and  
CC cytosstatic activity. The orally administered delayed-release form of  
CC phosphatidylcholine prevents premature resorption and provides for  
CC targeted release in the lower section of the small or large intestine.  
CC This sequence encodes the human MDR-3 protein described in the method of  
CC the invention.

Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Query Match	36.0%;	Score 1027.2;	DB 21;	Length 3924;
Best Local Similarity	61.9%;	Pred. No. 1.5e-262;		
Matches 1682;	Conservative	0;	Mismatches 1003;	Indels 33;
				Gaps 2;

OY	166	attatctgataaagaacccagatataagtaactctttccacagcctggtatataaacctgtaacc	225
Db	1140	attctctgataataataccttaaaatttgccagctttcttcagagagagacacaaacagacgc	1199
OY	226	atagaagaactctgtagatttaaaatgtcttcttcaattacatccataccacatctac	285
Db	1200	atcaaaagggaatttgtagcttccaatgattgttcacttcttcaacctctccggctcaagtc	1259
OY	286	aagatctgaaaggtctgaaatctagaatlaagtcctgagagacagtcgcctgtgcgt	345
Db	1260	aagatcttgaaaggccctcaaccctgaaagctgacagatgaggagacagctgctcctgtttga	1319
OY	346	ctcaatgtagcagttgggaagaaatagctgtagctccagctctctgcagaagttatatgaccgat	405
Db	1320	agctgctgctctgttgaaagagcaacaacggtccacgctctgatacagaagctctcaagacctgat	1379
OY	406	gattgctcttatacagtggtgagatgagaatgcacatcaagagcttctaaagtctccgcatatcga	465
Db	1380	gagggcacaaatcaactctgtctggcgagatattaggaacttaatgataacatctgag	1439
OY	466	gaccatattgtagtctgcttagtccaagagcctgtttcttcgggacacacatcagtaacat	525
Db	1440	gaatcatctgtgtgtgtagatccagagccggtctcttttccacacacatctgctgaat	1499
OY	526	atcaagaatgtagcagagatctgtctgactcgatgaaagatggaagagcagcaagaggaaga	585
Db	1500	attgttatgtagcgtgtgaagaatgtacacatgagatgagataagaagaagcctgtcaagaagcc	1559
OY	586	aatgctatgatttcatcatgagagttctcctaataaattatatacatctgttagggagaaa	645
Db	1560	aagcctctatggtttatcatcagaatattaccacagaaaatttgacacacctgttgagagaga	1619
OY	646	ggaagctcaatgagttagaaggcgagaaacaaagagatctgcattgctcgtagcttaattga	705
Db	1620	gggggccagcgtgagtggtggcgagaaacaaagagatctgcattgcaagtgacctgtgttcgc	1679
OY	706	aaccccaagatctcgaatttllagaatggagctcagctctgcacctgattcagcaagaagctca	765
Db	1680	aaccccaagatcctctctgctgtagtgagggccaagctcagcatgtgacacagaaagtgaagct	1739
OY	766	gctcttcaagctcagctagsgagaaagcgagcaaaagctcagactacaatcgtgtgtagcacac	825
Db	1740	gagtagcagcgagctctgtgataagggccagagaaagcgagacacacattgtgtgataattc	1799
OY	826	cgactcttactatctgaagttagcagattgttagtgtagccctaaagatgtgaattgttgcg	885
Db	1800	cgagcttcaagcggtccggaatgcagagatcatcgtcgtggtttggaagatgagataattctg	1859
OY	886	gagaaagagcacatctgctcgaactaatgycanaacgggctataatattcacatttgtatg	944
Db	1860	gagcagaaggaagccacagcgaactgtatgagaagaagggggtgtactctcaacttgtcaac	1919
OY	945	-----gtcacagatatataaaagctgtatgacagatg	978
Db	1920	atgcagacatcaggaagccagatccagctcagaaagaaatttgactactaaatgtgaaaagcct	1979

QY	979	gafctacatgacatacttctacttgaaagaagaccacactcaattccctctgacactcgtgaaag	1030
Db	1980	gccactgaagatgagccccaatgctggaatctgcgcctattgaagcattctactccagaag	2039
QY	1039	agcatcaagtacagctctcatctgacaagagc-----tgaagatccaccaccaattcaagag	1092
Db	2040	aaccttaaaattcaacaatgctgcgaagagagccttgatgctggaaccgagtgagcttgaa	2099
QY	1093	ataagctccctgagagctctctctataaanaatttaagaatttaacaagacgtgaatgcct	1152
Db	2100	gcaaatgycacacagagctgcctctctctgaaagctgcctcgaactgaataaacaagaatgccc	2159
QY	1153	tttggtgctcggggagacattgctctctgtctcaaatgagacgtgcacacagatatttcc	1212
Db	2160	tacttgcgtcgtggaacagatagctgcacatgcccgaatggggcttcacgcggacttcca	2219
QY	1213	atcacctcttgcaaaaatcataaccagctgttggaataatgataaataaacaacataaagcat	1272
Db	2220	gtcatattctcaagatcatcaagcatatttttggaaccagsgagatgacagctgaagcagcat	2279
QY	1273	gatgcaanaatttcatccatgatatcgcctaatcttgggtgtgattttgcttgcagtta	1332
Db	2280	aagtgcaacatattctcttgattcttctattctctgagattattctcttttacttcc	2339
QY	1333	ttcatgcaaggagatatttttcgcgcgagagcaggggaaattttaagatgagattagaac	1392
Db	2340	ttcctctcagggcttccacgcttcttggaagagctctgagatcctccacagaagactgcggtca	2399
QY	1393	ttgsgcttcaagacagctgttataccagatagctgcgtgttggaagaaagaaagacgc	1452
Db	2400	atgagcttttaagaagacatgacaaagacaggaatgagctgtgttgatgaccataaanaacgt	2459
QY	1453	acagagagctggaacaacaatattagccatagatatagcacaatccaaatccagaagcagagt	1512
Db	2460	actggtgacacttctacaagaacttgcacagatgctgcacgaagccagaagccacaagga	2519
QY	1513	ttccagatctgcgctcttaacacaaaatgcaactacaatgagagcttccagttacatttcc	1572
Db	2520	accaggtctggtcttaattctgacagaatataagcttcaaccttggaactcgtgtatatcatca	2579
QY	1573	tttataatgataggggagatgacatccctgatactctgagatctgagctccagactcctgcgtg	1632
Db	2580	tttatctacggtctgacgtttaaccctatgtcttatagcagctgttccaatatctgcgtg	2639
QY	1633	acaagatgattgaaacccgcagcagcatgctgatatctgccaanaaagaatgaagcaagaact	1692
Db	2640	tcaaggaattgctgaatgaaattgttgctctggaatgccaagaagatataaanaagaactg	2699
QY	1693	aagcatctctggaagaatagcgaactctgaaagcttggagaaatatacgtacatagtgtcat	1752
Db	2700	gaagctctctggaagaatctgcaacagagcaatatagaataattatlagacagctgtgctctg	2759
QY	1753	acaagggaaaagagccttcgcgaacatgtatgaaagagatgcttccagatccaaocacagaat	1812
Db	2760	accagaagaaagaaaattgtaatcaatgtatgttgaanaattgtatgagacttaacagaat	2819
QY	1813	acctcgaagaagaacagatattgtagagctgttatgatactcaagcagcttataat	1872
Db	2820	ctctgctcgaagaagcacacatctatagagattactttagtatctcaagaacattatgatat	2879
QY	1873	tttgcctatgacagaggttcgatttggaagcctatttaatacgaagctgagcagaatgac	1932
Db	2880	tttctctatgcggtgtgttttgattgttgctgcatatccatcatgtgatatgacatatgagc	2939
QY	1933	ccagaagggcagctcatagtttttctactgcgaattgcatatvgagctatgagcactcgga	1992
Db	2940	ttcaagaagatgtatctcgtgttttctctgcgaattgttatcttggtgcaggtgctgaagac	2999
QY	1993	atgcgccttttggtgcctcgaatctccaagaacaaatcggggctgcgcagctatgttgc	2052
Db	3000	gccagttcaatttgcctccagactatgacaaagctaaagctgtctgcagccacttatctatg	3059
QY	2053	ttgttggaagaagaaaccaaatatatagacgcgcgagctcaagaagggganaaagccagaca	2112

```

Db 3060 ctgtttgaagaacaacccctgattgacagctacagtagaagggtctgaagctgtataaa 3119
OY 2113 tctgaagggaatttagagttctcgagaagtcctcttctctatccatgctgcgccaagtgt 2172
Db 3120 ttctgaagggaattataacatttaagtctgttctcaactatccaccgagcaaacgtg 3179
OY 2173 ttcatccctcggtgttaccctcagttatgagcgagaagacagtgatcttggg 2232
Db 3180 ccagtgcttcagggtgagctgaggtgtagaagaagccagacactagccctgtgtggc 3239
OY 2233 agcaagcgctgtggaagaagcactctgttcaactctgcagagacttatgacctgt 2292
Db 3240 agcagtgctgtggaagaagcagcaggtgtgccagctccctggagcggttctacgacctgt 3299
OY 2293 caagagcaagtgtgttctgaagtgtgtagatgcaaaagaattgaatgtacagtgtccgt 2352
Db 3300 gcgggagacagtgcttctcgtatgtgtcaagaagaacatcaatgtccagtgctcaga 3359
OY 2353 tcccaaatagcaatcgttccctcagaagcctgtgtcttcaactatgacattgtcagaac 2412
Db 3360 gctcaactcgaatcgtgtctcagaagcctacattgtactgtacagatgtccgagaat 3419
OY 2413 atcgctatgtgtaacaagccgtgtgtgtccattagatgagatcaagaagccgcaat 2472
Db 3420 attgctcttgagagaacaagccggtgtgtatcacagatgaaatgtgtagtccagcaaa 3479
OY 2473 gcaagcaaatccattcttatttgaaggctccctgagaatatacaacaacaagttgga 2532
Db 3480 gctgcacaatcatcctcttctcagagcgttaccaccaaaatgataagaagttggga 3539
OY 2533 ctgaaaggagcagcacttctgtgagcggaagaagaagactagctattgcaaggctctt 2592
Db 3540 gataaaggagacccaagctccagaaggtcacaagaagatgtctatgtcccgagccccc 3599
OY 2593 ctccaanaacaaatttattgtgtgtagagcgacacttcagccctcagataatagacgt 2652
Db 3600 atcagacaaccccaatcctcctctgtgtgtagagctacatccgctcgtgatatcgaagt 3659
OY 2653 gagaaggtgtgtcagcatgtccctgtataaagccagagcaggaagacatgtcctagtgtc 2712
Db 3660 gaaaggtgtgtcacaagaagccctgtgacaagaagagcgacacgtcgtatgtgatt 3719
OY 2713 actcagaagctctgtcgaatttcaagaagcagattgtagtgtcttgcacaaatggaaag 2772
Db 3720 gctcacgcctgtccacatccacatcagaatgcagacttaagtgtgttcaagaatgggaga 3779
OY 2773 ataaaggacaagaagacatcatcaagaagctcctgagaatcagagacatatatttaagta 2832
Db 3780 gtcaaggagcatgtgacagcatcagcgtcgtgcacagaagaagcattattttccaatg 3839
OY 2833 gtgaatgcacagtcagt 2850
Db 3840 gtccagtgctccaggtctgg 3857

```

## RESULT 15

AA004522 standard; DNA; 4378 BP.

AA004522;

XX 01-OCT-1990 (first entry)

XX Multidrug Resistance A1 gene.

XX Multidrug Resistance A1 gene; drug resistance of cancer cells; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH 138..3981

FT CDS /\*tag= a

```

XX JP02100680-A.
XX 12-APR-1990.
XX 05-OCT-1988; 88JP-0251475.
XX 05-OCT-1988; 88JO-0251475.
XX (SUNR ) SUNTORX LTD.
XX WPI: 1990-159707/21.
XX P-PSDB: AAR04868.
XX MDR related gene derived from human normal cells -
XX codes specific amino acid sequence, used for diagnosing drug
XX resistance of cancer cells
XX Disclosure: ; P; Japanese.
XX The gene is useful for diagnosis of drug resistance of cancer
XX cells.
SQ Sequence 4378 BP; 1313 A; 819 C; 1058 G; 1188 T; 0 other;

Query Match 36.0%; Score 1027; DB 11; Length 4378;
Best Local Similarity 62.1%; Pred. No. 1.8e-262;
Matches 1693; Conservative 0; Mismatches 990; Indels 42; Gaps 3;

OY 166 attattgataaagaacccagatagataactttccacagctcgatataaactgaatcc 225
Db 1239 ataattgataaataaagccaagtatttgcagctatttgcgaaggtggaacaaacagataat 1298
OY 226 atagaaggagactgttgaaatttaaaaatgttcttcaattatccatccaagaccatc 285
Db 1299 attaaaggaaatttgaaattcagaagaatgttcaacttcagctaccatccgaaagaagt 1358
OY 286 aagattctgaaaggtcttgaaatcagaatcgaatgaatcgtgagagaacagtcgctgtgtc 345
Db 1359 aggaattctgaaaggtcttgaaatcagaatcgaatgagagtgagagacggtggtcgtgtgga 1418
OY 346 cccaatgtcagtgaggaaggtacggttagtcagccttccagaggttaatgatccgagt 405
Db 1419 aacagtgctgtggaagaagcacaacagtcacgctgatatcagaagctctcatgacccaca 1478
OY 406 gatgcttatcatcgtgtgtagatgaatgaatcagaagcctttaaattgtgcgcatatcga 465
Db 1479 gagggagatgtgtgtagatgtagacaggaatttagacataaattgaagtttctcagt 1538
OY 466 gacatattgaggtgtgttagtcaagaagcctgttgttctgggacccatcagtaaat 525
Db 1539 gaatcattgtgtgtgtgtagacgaacctgtattgttgcacacagatagctgaaac 1598
OY 526 atcaagtatgagcagagatgtgtacgtgaagaagatgagagcagaacgaagcaaggaagca 585
Db 1599 attcgtatggtccgtggaataatgcacatgagtatgagatgagaagaagcttcaagaagc 1658
OY 586 aatgctatgatttattcatcgtgagttctcctaataaatttaatacatgtgttaggggaaaaa 645
Db 1659 aatgctatgatttattcatcgtgagttctcctaataaatttaatacatgtgttaggggaaaaa 1718
OY 646 gtagctcaaatgagtggaagcagaagaacagagatgcgaattgtcgtgtccttagtctga 705
Db 1719 ggggccagttgagtgtgtgtagaagcagaagatgcagatgtcagcgtgtccctgtgtcgt 1778
OY 706 aacccaagaattctgatttagatagagctgacgtgtcgtcgtgattcgaagaagcaagta 765
Db 1779 aacccaagaattcctcgtctgtatgtaggccaagtcagcgtcttgagacacagaagcgagaaga 1838
OY 766 gctgttcaagctgtcagtgagaagcgcgagcaaaagtcggaactacaatcgtgtgtagacac 825
Db 1839 gtgttctcagtggtccttgataagcgacagaagaagtcggaacacatctgtgtatgctcat 1898

```

OY	826	cgactcttactatcttcgaagctgcagatttgatttgcagccctaaaggatgtaagctgcggc	885
Db	1899	cgcttgcctacagcttcctgattctgtaacgcatccgcgtgcttcgattgtaaggatcatcttg	1958
OY	886	gagaaagagcgacatgcgtcgaactatgtgcaaacgagagcgtctatattcatctgcgtac	945
Db	1959	ggaaaaaggaatacctcatgtaactcatgaaaggaaggaagcatttactctcaactctgcaca	2018
OY	946	tcaasgatatataaaaagctgcatacagatggaagtcacatattcta-----	997
Db	2019	atcgagacagcaggaatatgaattgaataaaatgcagcgtgatgatccataaagtga	2078
OY	998	-----ctbaaagaagaacaaactccctccctgcactcgtgcactctggaagagatca---	1045
Db	2079	attgatgccttgaaatgctctcccaaatgatccaagctcaagcttaataaagaatca	2138
OY	1046	-----agtcagacttcattgacaagcgtgagaaatccaccacactc	1086
Db	2139	actcgtaggagtgctccgtgatacacaagcccaagacaagaagcttagtaccaaagagct	2198
OY	1087	aaagagataagcttccttcgaagctctctcatttaaaattttaagttaaacaagcctgta	1146
Db	2199	ctgatagaagaatataccctccagcttcctcttggaagattatgaagctaaatttaactgaa	2258
OY	1147	tggcctttgttgcttcgggaacatgcgtctcgtcttaattgaaactgcttcacgtta	1206
Db	2259	tggcctattttgtttgttgatcttcttgccattataaavgagcgctgcgaaccgca	2318
OY	1207	ttcttcacatcctctgcataaaatataacacatgcttggaataatgataaaacc---aca	1263
Db	2319	tttgcacatatattcttcaagaattacagaaggtttcttaacaagaattgatgcttcgaaca	2378
OY	1264	ttaagcagatgcagaaattttccacatgatatctgcactttgggtgatctgctt	1323
Db	2379	aaacgacgaatagtactgtttccacattgtttccagccttggaattatttctt	2438
OY	1324	gtcagttatttcaatcagaggaattatttaccgcagagcaggggaattttaacgatgaa	1383
Db	2439	attacattttccctcagggcttcacatttgcagaagctggaagatccctcacacaagcgg	2498
OY	1384	ttaagcacttgccttcaagcactgtttatccaggaatttcgcgtgttgattgaaag	1443
Db	2499	cttcgataacatggttttccgataccacatgcctcagacaggaatgvtggttgattgaaccc	2558
OY	1444	gaaacaacgacagagagcttgcacaacataattgcacatagataatgacacaattcaagaa	1503
Db	2559	aaaaacacacacttgagcatgtaccagaagctgcgcaatgatgtcgtccaagttaaagg	2618
OY	1504	gcaaacggttccaggaattggtcgtcttaacacaatgcaactcaactatgggaacttccgt	1563
Db	2619	gcatatggttccagggcttgcgtgaattacccaagaataatgaacaaactcttgcagagata	2678
OY	1564	atcatctcccttatatgtatgtggaatgacatcttcctgattctgcagttatgtcccgta	1623
Db	2679	atatataccttaactatggttgcgaactcaacactgttactcttagaatgttgaccactc	2738
OY	1624	cttgcgcgtgcaggaatgatattgaaaccgcagcaatgatcgtgatttgcacaagaataag	1683
Db	2739	attgcacaaatagcagagtgcttctgaatbanaaattgtctgcgaacagacatgaaagaatag	2798
OY	1684	caaggaacttaagcatgcttgcgaagatgcacacatgaagccttgcggaataatagctacata	1743
Db	2799	aaagaactagaaaggtcttcgggaagatcgcctacatgaagcaatagaanaactccgcgaacgct	2858
OY	1744	gtctcattaaacaaggaanaaagccttcgaagcaaatgtatgaagaagcttcaagactcaa	1803
Db	2859	gttctcttgctcagagcgaggaagtgttgaacatatgtatgctcagagtttgcaggtacaa	2918
OY	1804	cacagaataatccttcogaagaagcacagatatcttgcgaagcgtgtatgtcatatgaaccagcc	1863
Db	2919	tacagaacactcttgcgaagaagcacacactcttgcgaataatttccctacaccagcga	2978

QY	1864	lctalatalatttgcctacagcagcagggttcttgatttggagccatttaattcaagctcgga	1923
Db	2979	atgagtatcttttccctacatctctgagatgtttcccggttttggagcctactcgtgtgagcacaataa	3038
QY	1924	cgaatgagcccccagaagggcagtgctcatcaagtttttaactgaatttcgatatatggagctatgccc	1983
Db	3039	ctcatagagccttttgaagatgttctgtttagatttttaagctgtgttgcctttgtgtgcacatgccc	3098
QY	1984	atcggaaaaaacgctcgttlttggctctctgaataattcccaagccaaatccggggtctgcgat	2043
Db	3099	gtggggagaagtcagttcaatttgccttgcgactatgccaagacaaataatcacagcggccac	3158
QY	2044	ctgttttccttgttggaaaaagaacccaatatagacagccgagtcgaagccaaggaanaag	2103
Db	3159	atcatcatgatactgaaaaaacccctttgatttgcagctaaagcagaagagcctaatg	3218
QY	2104	ccaacacatgataagaggaattttagagtttctgagagagctctcttcttccatccatgtcgc	2163
Db	3219	ccgaacacattggaaggaataatgtcacattttgttggagttgttatccaactccaccgga	3278
QY	2164	ccagaatgttttcatcctcgttgcgttactatccctccagtatattgagcgaggaanaacagttgca	2223
Db	3279	ccggagacatcccaatgtcttccagtgagcagctcgagagtgaaagaagygccaaacgctggtc	3338
QY	2224	tttttggggagcagcggcctcttgggaaaaagcaactctcgttccaactctctcagaagacttat	2283
Db	3339	ctgtgttgagcagcagctgtgccttctgtgggaagagacagatggtgcagctctccgttgaagcttcc	3398
QY	2284	gaaccccgctgcagaagacaaagtctcgttttggatgtgtgagatgcanaagaattgaatgtcacg	2343
Db	3399	gaaccccttggagaggaagaagatgctcgtcttggatgtggcaagaataaagcagctgaatgttccag	3458
QY	2344	tggtcttcggttcccaataagacaactcgttctccaagagccctgtccttccaattgagcatt	2403
Db	3459	tggtcttcggaagacacccctgggacacgtgctccccaagagcccaatccctgttgcattgcagcatt	3518
QY	2404	gctgagaacaatcgccctctatgtgtgacaacagccggtgtgtgtgcattagatlatgatacgaaga	2463
Db	3519	gctgagaacaatctgctctatgttgagacaacagccgggtgtgtgttcaacaggaagagattgttgag	3578
QY	2464	gccgcgaatgtgcagaataatcatcattctttatttgaagagctctccctcgagaataatacacaaca	2523
Db	3579	gcaagcaaaaggggccaacatacatatgccttccattatcgaagctctctaatanaatatagcact	3638
QY	2524	caagtttgcacctgaaagagagacagagcttcttcggcgccagaaaaaagaagcctagcttatgca	2583
Db	3639	aaagtaagagacaagaaggaacccagcctctccgtgtgtgcccgaagaaacagcatgtgcacatgct	3698
QY	2584	aggcctctctccccaanaaacccaataatttatgttggatgagagcaacttcaagccctcgat	2643
Db	3699	cggtgcctctgtttagacagagccctcatatatttgccttttggatgagagcaagcagctcgtgat	3758
QY	2644	aatgacagttgagaaggtgtgtctcagcatagccctctgtataaagccaagagcggaagagacatgc	2703
Db	3759	acagaagaagtgtgaaaaggtgttctccaagaagccccttgagacaagaaggaagggccgcgcgcgc	3818
QY	2704	ctagatgggtcactccacagagctctctgcacattccagaacagcaggaatttgaatgtgttctgac	2763
Db	3819	attgtgattgtctccgcgcgtctccacacatccagaaatagcagacttaataagttggttttcag	3878
QY	2764	aatgaaagataaagggacaaggaactcatccaagagctctcttgaagaatccgagacatatat	2823
Db	3879	aatggtcagagtcacaaagagagcatgtgacagcatcagcagagctgtctgtgcacagaagaagcatctat	3938
QY	2824	tttaagttagtgaatgcacagttcag	2848
Db	3939	tttccaatgtcagttgtccaggtcg	3963

Search completed: April 22, 2002, 23:02:05  
Job time: 25713 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:31:37 ; Search time 9392.76 Seconds  
(without alignments)  
3267.404 Million cell updates/sec

Title: US-09-873-409-10

Perfect score: 2856

Sequence: 1 cctaatccctaatatc.....atgcacagtcagtcagtga 2856

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpl:\*  
6: em\_estba:\*  
7: em\_estrov:\*  
8: em\_estov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_prod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	545	19.1	545	10	AL040762 DKF2P434C
2	386.6	13.5	943	10	AL520322 AL520322
3	357.4	12.5	559	11	BF692596 602248949
4	279.6	9.8	405	10	AA243820 AA243820
5	275.2	9.6	726	11	BG293345 602390738
6	273.2	9.6	1019	11	BG248052 602359987
7	270.8	9.5	894	11	BF584668 602098406
8	269	9.4	944	11	BF796582 602258463
9	262.2	9.2	785	10	AV709991 AV709991
10	252.4	8.8	560	10	BE749379 200194 MA
11	249.4	8.7	981	11	BF696667 602272046
12	249	8.7	795	11	BF313560 601500192

13	248.8	8.7	886	13	BH139685
14	242.6 <td>8.5 <td>780 <td>11 <td>BG587938</td> </td></td></td>	8.5 <td>780 <td>11 <td>BG587938</td> </td></td>	780 <td>11 <td>BG587938</td> </td>	11 <td>BG587938</td>	BG587938
15	241.4 <td>8.5 <td>886 <td>13 <td>AE540627</td> </td></td></td>	8.5 <td>886 <td>13 <td>AE540627</td> </td></td>	886 <td>13 <td>AE540627</td> </td>	13 <td>AE540627</td>	AE540627
16	239.6 <td>8.4 <td>897 <td>13 <td>AE541090</td> </td></td></td>	8.4 <td>897 <td>13 <td>AE541090</td> </td></td>	897 <td>13 <td>AE541090</td> </td>	13 <td>AE541090</td>	AE541090
17	237.4 <td>8.3 <td>853 <td>13 <td>AE679807</td> </td></td></td>	8.3 <td>853 <td>13 <td>AE679807</td> </td></td>	853 <td>13 <td>AE679807</td> </td>	13 <td>AE679807</td>	AE679807
18	237.4 <td>8.3 <td>880 <td>13 <td>AE687805</td> </td></td></td>	8.3 <td>880 <td>13 <td>AE687805</td> </td></td>	880 <td>13 <td>AE687805</td> </td>	13 <td>AE687805</td>	AE687805
19	237.4 <td>8.3 <td>932 <td>13 <td>AE670821</td> </td></td></td>	8.3 <td>932 <td>13 <td>AE670821</td> </td></td>	932 <td>13 <td>AE670821</td> </td>	13 <td>AE670821</td>	AE670821
20	236 <td>8.1 <td>834 <td>13 <td>AE548312</td> </td></td></td>	8.1 <td>834 <td>13 <td>AE548312</td> </td></td>	834 <td>13 <td>AE548312</td> </td>	13 <td>AE548312</td>	AE548312
21	230.6 <td>8.1 <td>871 <td>13 <td>AE682350</td> </td></td></td>	8.1 <td>871 <td>13 <td>AE682350</td> </td></td>	871 <td>13 <td>AE682350</td> </td>	13 <td>AE682350</td>	AE682350
22	230.2 <td>8.1 <td>823 <td>13 <td>AE532602</td> </td></td></td>	8.1 <td>823 <td>13 <td>AE532602</td> </td></td>	823 <td>13 <td>AE532602</td> </td>	13 <td>AE532602</td>	AE532602
23	229.4 <td>8.0 <td>482 <td>11 <td>BI344244</td> </td></td></td>	8.0 <td>482 <td>11 <td>BI344244</td> </td></td>	482 <td>11 <td>BI344244</td> </td>	11 <td>BI344244</td>	BI344244
24	227.8 <td>8.0 <td>651 <td>10 <td>BE016688</td> </td></td></td>	8.0 <td>651 <td>10 <td>BE016688</td> </td></td>	651 <td>10 <td>BE016688</td> </td>	10 <td>BE016688</td>	BE016688
25	227.2 <td>8.0 <td>947 <td>13 <td>AE683755</td> </td></td></td>	8.0 <td>947 <td>13 <td>AE683755</td> </td></td>	947 <td>13 <td>AE683755</td> </td>	13 <td>AE683755</td>	AE683755
26	226 <td>7.9 <td>687 <td>11 <td>BG646725</td> </td></td></td>	7.9 <td>687 <td>11 <td>BG646725</td> </td></td>	687 <td>11 <td>BG646725</td> </td>	11 <td>BG646725</td>	BG646725
27	225 <td>7.9 <td>919 <td>13 <td>AE690701</td> </td></td></td>	7.9 <td>919 <td>13 <td>AE690701</td> </td></td>	919 <td>13 <td>AE690701</td> </td>	13 <td>AE690701</td>	AE690701
28	224.2 <td>7.9 <td>842 <td>13 <td>AE671925</td> </td></td></td>	7.9 <td>842 <td>13 <td>AE671925</td> </td></td>	842 <td>13 <td>AE671925</td> </td>	13 <td>AE671925</td>	AE671925
29	216.4 <td>7.6 <td>891 <td>13 <td>AE682250</td> </td></td></td>	7.6 <td>891 <td>13 <td>AE682250</td> </td></td>	891 <td>13 <td>AE682250</td> </td>	13 <td>AE682250</td>	AE682250
30	215.2 <td>7.5 <td>899 <td>13 <td>AE548799</td> </td></td></td>	7.5 <td>899 <td>13 <td>AE548799</td> </td></td>	899 <td>13 <td>AE548799</td> </td>	13 <td>AE548799</td>	AE548799
31	214.4 <td>7.5 <td>921 <td>13 <td>AE687628</td> </td></td></td>	7.5 <td>921 <td>13 <td>AE687628</td> </td></td>	921 <td>13 <td>AE687628</td> </td>	13 <td>AE687628</td>	AE687628
32	213 <td>7.5 <td>519 <td>10 <td>AA434959</td> </td></td></td>	7.5 <td>519 <td>10 <td>AA434959</td> </td></td>	519 <td>10 <td>AA434959</td> </td>	10 <td>AA434959</td>	AA434959
33	212.2 <td>7.4 <td>605 <td>10 <td>AV616675</td> </td></td></td>	7.4 <td>605 <td>10 <td>AV616675</td> </td></td>	605 <td>10 <td>AV616675</td> </td>	10 <td>AV616675</td>	AV616675
34	210.4 <td>7.4 <td>750 <td>11 <td>BG585786</td> </td></td></td>	7.4 <td>750 <td>11 <td>BG585786</td> </td></td>	750 <td>11 <td>BG585786</td> </td>	11 <td>BG585786</td>	BG585786
35	208 <td>7.3 <td>851 <td>11 <td>BF696062</td> </td></td></td>	7.3 <td>851 <td>11 <td>BF696062</td> </td></td>	851 <td>11 <td>BF696062</td> </td>	11 <td>BF696062</td>	BF696062
36	203.8 <td>7.1 <td>661 <td>11 <td>BG302675</td> </td></td></td>	7.1 <td>661 <td>11 <td>BG302675</td> </td></td>	661 <td>11 <td>BG302675</td> </td>	11 <td>BG302675</td>	BG302675
37	201.6 <td>7.1 <td>575 <td>10 <td>AW342644</td> </td></td></td>	7.1 <td>575 <td>10 <td>AW342644</td> </td></td>	575 <td>10 <td>AW342644</td> </td>	10 <td>AW342644</td>	AW342644
38	200.8 <td>7.0 <td>402 <td>11 <td>H82550</td> </td></td></td>	7.0 <td>402 <td>11 <td>H82550</td> </td></td>	402 <td>11 <td>H82550</td> </td>	11 <td>H82550</td>	H82550
39	200.4 <td>7.0 <td>787 <td>11 <td>BG584063</td> </td></td></td>	7.0 <td>787 <td>11 <td>BG584063</td> </td></td>	787 <td>11 <td>BG584063</td> </td>	11 <td>BG584063</td>	BG584063
40	199 <td>7.0 <td>854 <td>13 <td>AE682568</td> </td></td></td>	7.0 <td>854 <td>13 <td>AE682568</td> </td></td>	854 <td>13 <td>AE682568</td> </td>	13 <td>AE682568</td>	AE682568
41	197 <td>6.9 <td>541 <td>10 <td>AI722858</td> </td></td></td>	6.9 <td>541 <td>10 <td>AI722858</td> </td></td>	541 <td>10 <td>AI722858</td> </td>	10 <td>AI722858</td>	AI722858
42	196.4 <td>6.9 <td>2676 <td>12 <td>AK014319</td> </td></td></td>	6.9 <td>2676 <td>12 <td>AK014319</td> </td></td>	2676 <td>12 <td>AK014319</td> </td>	12 <td>AK014319</td>	AK014319
43	195.6 <td>6.8 <td>704</td> <td>10 <td>BE239652</td> </td></td>	6.8 <td>704</td> <td>10 <td>BE239652</td> </td>	704	10 <td>BE239652</td>	BE239652
44	193.8 <td>6.8 <td>815 <td>11 <td>BI409057</td> </td></td></td>	6.8 <td>815 <td>11 <td>BI409057</td> </td></td>	815 <td>11 <td>BI409057</td> </td>	11 <td>BI409057</td>	BI409057
45	192.6 <td>6.7 <td>511 <td>10 <td>AA239727</td> </td></td></td>	6.7 <td>511 <td>10 <td>AA239727</td> </td></td>	511 <td>10 <td>AA239727</td> </td>	10 <td>AA239727</td>	AA239727

#### ALIGNMENTS

RESULT 1  
LOCUS AL040762 545 bp mRNA EST 29-FEB-2000  
DEFINITION DKF2P434C1815.F1 434 (synonym: htes3) Homo sapiens cDNA clone  
KEYWORDS DKF2P434C1815 5', mRNA sequence.  
ACCESSION AL040762  
VERSION AL040762.1 GI:5409708  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 545)  
AUTHORS Bloecher,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
TITLE EST (Bloecher, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Bloecher H

MIPS  
Am Kioferpitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,  
Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
s1 sequence also available.  
This clone (DKF2P434C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubneweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

#### FEATURES

source  
1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKF2P434C1815"



```

/clone_lib="434 (synonym: hres3)"
/issne_type="testis"
/dev_stage="adult"
/lab_pos="DH10B"
/mote="vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a      89 c      138 g      147 t
ORIGIN

```

Query Match	Similarity	19.1%	Score 545:	DB 104:	Length 545:
Best Local	Similarity	100.0%	Pred. No. 8.9e-123:		
Matches	545:	Conservative	0:	Mismatches	0:
				Indels	0:
				Gaps	0:
OY	182	ccagatagataactcttcacacagcttgatataaaacctgaaccatcagaaggaaactgttg	241		
Db	1	CCAGATATGATTAACCTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAAGAACTGTGG	60		
OY	242	aatttaaaaatgtcttcttcaattatccatcaagaaccatctatcaagaattctgaaagtc	301		
Db	61	AATTTAAAAAATGTTCTTTCAATATTCATCAAGACCAATCTATCAAGATTCTGAAAGGTC	120		
OY	302	tgaactcagaatgaatgaatcgtgagagaaagctgcgcttggtgtgttcctcaatgacagtgga	361		
Db	121	TGAATCTCAAGATTAAGTCTGGAGAGACAGTGCCTTGCTGCTTCATAGCGCATGGGA	180		
OY	362	agagacggtatccacagctctcgcagaggtatgatccggatgatgagcttatcatg	421		
Db	181	AGAGTACGGTACTCCAGCTTTCGACAGGCTTATATGATTCGGATATGCTTTATCATGG	240		
OY	422	tggatgagaatpacatcagaagctttaaattgycggagcatatcgagaaccatattggaatg	481		
Db	241	TGGATGAGATGACATCAAGAGCTTTAAATAGTGGCGCATATTCGAGACCATATTGAGATGG	300		
OY	482	ttagtcagaagccgtgtttgttcgggaccacacatcagtaacatatcaagatgatgacgag	541		
Db	301	TTTAGTCAAGAGCGCTGTTTGTTCGGGACCCACCATCAGTAAACAATATCAAGTATGGACGAG	360		
OY	542	atgagtgaactatgaagagatgagagagagcagaaggaaacaaatgctatgattta	601		
Db	361	ATGATGTGCTATGAGAGATGGAGAGAGCAAGCAAGGGAAGCAAAATGCTATGATTTA	420		
OY	602	tcaatgagcttccataaattatactatgtaaggggaaaaagagatcacaatagtg	661		
Db	421	TCATGGAGTTCCTATAAATTAAATTAACTTGGTAGGGGAAAAAGAGCTCAAAATGAGTG	480		
OY	662	gagggcagaagaagagatcgcaattgctcgtgacttagtcgnaaaccacaagatctga	721		
Db	481	GAGGGCAGAAACAGAGGATCGCAATTGCTGCTGCTTACTTGAAACCCCAAGATTCTGA	540		
OY	722	tttta 726			
Db	541	TTTTT 545			
RESULT	2				
LOCUS	AL520322	943 bp	mRNA	EST	13-FEB-2001
DEFINITION	AL520322 LTI_NFL004_NBC3 Homo sapiens cDNA clone CS0DB006YCI5 5				
ACCESSION	AL520322				
VERSION	AL520322.1	GI:12783815			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.				
TITLE	1 (bases 1 to 943)				
JOURNAL	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
COMMENT	Full-length cDNA libraries and normalization				
	Unpublished (2001)				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				

Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).  
 Location/Qualifiers  
 1. 943  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB006YC15"  
 /clone\_lib="LFI\_NFL004\_NBC2"  
 /sex="male"  
 /issue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [liangellife@life.com](mailto:liangellife@life.com) URL : <http://fulllength.invitrogen.com>"  
 BASE COUNT 253 a 225 c 253 g 211 t 1 others  
 ORIGIN  
 Query Match 13.5%; Score 386.6; DB 10; Length 943;  
 Best Local Similarity 65.0%; Pred. No. 6.1e-84;  
 Matches 586; Conservative 0; Mismatches 314; Indels 1; Gaps 1;  
 Oy 1949 tagtttactcgaatgcataatgagatgatggccatgcgaaaaagctcgttttgctc 2008  
 Db 6 TATATTTTTCAGCGCTGTGCTTGGTGCCATGCGCGTGGGCAAGTCAGTTCATTGCTC 65  
 Oy 2009 ctgaatattccaagccaatcgggggctgcgcatactgtttgcttgttggaaagaac 2068  
 Db 66 CTGACTATGCCAAACCCAAATATGACGAGCCCATCATCATGATCATTTGAAAAACC 125  
 Oy 2069 caaatatgacacgcgcgcgactcaagaaggaaaaagccaagaacagatgaaggaaattag 2128  
 Db 126 CTTTGATTGACAGCTACAGCAGCAGGAGCCCTATATGCCAACAACATTGGAGGAATGTCA 185  
 Oy 2129 agttcgaagaatcctctctctctatcatcatgctgcgcagaatgtttcatcctccgtgct 2188  
 Db 186 CATTGTGGTGAATGTTATTTCACTATCCACCCGACCGGACATCCCAATGCTTCAGGGAC 245  
 Oy 2189 tatccctcaagtatgacgcgaggaagaacagatgaactgttggggagacgcggctgtggga 2248  
 Db 246 TGAGCCTGAGAGTGTAAGAGGCGCAGACGCTGGCTCTGCTGGTGAGCAGTGGCTGTGGGA 305  
 Oy 2249 aaagcactctgtttaaactcttgcggagaaactttaagaccgcgtgcaagaagaatgctgt 2308  
 Db 306 ACAGCACAGTGTCCAGCTCCTGAGGCGGTTCTACGACCCCTTGGCAAGGAAGTGTGC 365  
 Oy 2309 ttgatgtgtgtgatgcaaaaagaatgtaattgacatgtgctccgttcccaaatagcaatcg 2368  
 Db 366 TTGATGGGCAAAATAATTAAGCGACGATGTTCAATGGCTCCGACACACCTCGGGCAATCG 425  
 Oy 2369 ttccctcaagagcctgtgctcttcaactgcgaagcattgtctgagaacatcgccatagtgtgaca 2428  
 Db 426 TGTCCACAGAGCCATCCGTTTGACTCAGACATGTGTGAGAACATTCCTATGAGAGACA 485  
 Oy 2429 aacagccgtgtgtgcatatagatgataatcacaagaagccgcgaatgcagcaaatatccatt 2488  
 Db 486 ACAGCCGGGTGTTGCACAGGAGGATCTGTAGGGCAGCAAAAGGAGGCCAACATATCATG 545  
 Oy 2489 ctttatattgaagttccctctgagaataatcaaacacaagtttgagctgaagaagagcagc 2548  
 Db 546 CCTTATCGAGTCACTGCTTATTAATATATAGCATTAAAGTATGAGACAAAGGAACTCAGC 605  
 Oy 2549 ttcttcggcgccagaacaagaactcagctattgtcagaaggtctcttccaaaaacccaataa 2608  
 Db 606 TCTCTGTGTGGCCAGAAACAGCATTGCCATGTGCTGTGCTTGTAGACAGCCTCATTA 665  
 Oy 2609 tttaattgttgatgagggcacttcagccctcgataatgacatgacatgagaagtggttcagc 2668

Query Match	13.5%;	Score 386.6;	DB 10;	Length 943;
Best Local Similarity	65.0%;	Pred. No. 6;	le-84;	
Matches	586;	Conservative	0;	Mismatches 314; Indels 1; Gaps 1;

QY	1949	taagtttactgaatcgaatgcataatgagatcgtatgacatgcgaanaaagctcgttttgctc	2008
DB	6	TATATTTTTCACGCTGTGTCCTTTGGTGGCAGTGGCGGTGGGCAATCAGTTCAATTGCTGC	65
QY	2009	ctgaatattccaagaagccaatcggggctgcgcatactgtttgccttgtltgaaanaaac	2068
DB	66	CTGACTATGCGCCAAACCCAAATATCAGACGGCCCATATCATGTATCATTTGAAAAAACCC	125
QY	2069	caaatatgacgcgcgcagctcaagaaggaaaaagccagacacatgtaagaagattag	2128
DB	126	CTTTGATTGACAGCTACACGACGGAAGGCTCTATATCCCAACACATTGGAGAAATGTCA	185
QY	2129	agttccgaagaatccctctctctctatcacaatgctgcacagatgtttcatalccctcgtgct	2188
DB	186	CATTTGGTGAATTTATTTCAACTTCCACCCGACCGGACATCCCACTGCTTACGGGAC	245
QY	2189	tatccctcaagtatctgacgcgaggaagaacagatgaacttgtgtggagacgcggtctgtgga	2248
DB	246	TGAGGCTTGAGAGTGTAAGAAAGGCGCAGACGCTGGCTCTGCTGGCGAGCATGGCTGTGGGA	305
QY	2249	aaggaactctgtttaactcttgcgcgagaacttatagaccctgtgaagaagaagtgtgt	2308
DB	306	ACAGGACAGGTGTCACGCTCCTCGAGCGGTTTACGACCCCTTGGCAGGAAAGTGCTGC	365
QY	2309	tgtatgtgtgtgatatgcaaaaagaatgtaatgtacatgtgctccgttccccaatagcaatcg	2368
DB	366	TTGATGTGGCAAAATAAATAAGCGACGTGATGTTCATGTGGCTCCGACACACCTCGGGCAATCG	425
QY	2369	tctcctcaagagcctgtgctctcaactgcgaacatgtctgtgagaacatcgccatatgttga	2428
DB	426	TGTCCACGAGAGCCATCCCTGTTTGACTCGACGATGTGTGAGAACATTCCTATGAGAGCA	485
QY	2429	aagagcctgtgtgtgcattagatgtgatgccaaagaagccgcgaatgtcagcaaatatccatt	2488
DB	486	ACAGCCGGGTGTTGCACAGAGAAATGCTGAGGGCAGCAAAAGGAGGCCAACATATCATG	545
QY	2489	ctttatttgaagtcctccctgtgaataacacacacaaagtcagttgactgtgaagaagcagacgc	2548
DB	546	CCCTTATCATGACTGCTTATTAATATATAGCATTAAAGTATGAGACACAAAGGAACTCAGC	605
QY	2549	tcttcgcgcgcagagaacaaagactcagctattctgcagaaggtctcttccaaaacccaana	2608
DB	606	TCTCTGTGTGGCCAGAAACACGCAATTCGCATATGCTGCTGCTTTTGAAGACGCTCATTA	665
QY	2609	tttatatgttgtatgtaggcacttcagccctcgatataatgacatgtagaagatgtgttcagc	2668

```

Db      666  TTTTCTTTGGATGAGCAGCTCAGCTGATACAGAAAGTGTGTCAG 725
Oy      2669  atgcccctgtataaagccaggacggaagacatgctcctagtgtcactcaagcgtctctg 2728
Db      726   AAGCCCTGACAAACCCAGAGAACGCCACCTGCATTGTGTGATGCTCCACCCCTGTCCA 785
Oy      2729  caatcagaacgacagattgttagtg-gtctgcagaatggaagataaaggaaagga 2787
Db      786   CCATCCAGATGCACTTAATATGTTGTGTTTCAGAAATGCGAGATCAGAGCATGCGC 845
Oy      2788  actcctcaagaagcctcctggaatcgagacatatttttaagttagtgaaatgcacatga 2847
Db      846   ACGCATGACGACACTGCTGCGACAGAAAGCATCTATTTTCATAGTGTCCAGTCCAGCT 905
Oy      2848  g 2848
Db      906  g 906

RESULT 3
BF692596      559 bp      mRNA      EST      22-DEC-2000
LOCUS         602248949E1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
DEFINITION   mRNA sequence..
ACCESSION    BF692596
VERSION      BF692596.1 GI:11978004
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 559)
              NIH-MGC http://mgc.ncl.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: rgs@ncl.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              cDNA Library Preparation: CLONETECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://image.llnl.gov
              Plate: L10M1203 row: a column: 13
              High quality sequence stop: 555.
              Location/Qualifiers
                1..559
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="4334100"
                /clone_lib="NIH_MGC_62"
                /tissue_type="melanotic melanoma, high MDR"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
                SfiI (ggcgccctcgcc); Site_2: SfiI (ggccatlaagcc);
                Double-stranded cDNA was prepared from cell line RNA. 5'
                and 3' adaptors were used in cloning as follows: 5'
                adaptor sequence: 5'-CACGCCCATATGCGC-3' and 3' adaptor
                sequence: 5'-ATCTAGAGCGCGCGCGGCGGCGCATG-dt(30)-BN-3'
                (where B = A, C, or G and N = A, C, G, or T). Average
                insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                contained inserts by PCR. This library was enriched for
                full-length clones and was constructed by Clontech
                Laboratories (Palo Alto, CA)."
BASE COUNT   161 a 118 c 141 g 139 t
ORIGIN

```

```

Query Match      12.5%; Score 357.4; DB 11; Length 559;
Best Local Similarity 97.8%; Pred. NO. 7.9e-77;
Matches 394; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

```

```

Oy      394  tatgatccgagatgagccttatcatgtgtgatgaaatgacatcagagccttaatgtg 453
Db      1   TATGATCCCGATGATGCGCTTTATCAGGTGATGAGAAATGACATCAGACCTTTAAATGTG 60
Oy      454  cggacttacgagacacatatltgagtgtgtatgtaacaagaccgtttgtctggaccacc 513
Db      61   CGGCATTAACGAGACCATA-TGGAGTGTGTAGTCAAGACCGTG-TTGTGTCGGAGCCACC 118
Oy      514  atcagtaacaatalcaagatagagacagatgtgtactgtatgaaagagatgagagaga 573
Db      119  ATCAGTAACAATATCATAGATGATGAGAGATGATGATGATGATGATGATGATGATGATG 178
Oy      574  gcaagggaaagcaaatgagatgattatcatcagggatcttccaaataattataatgtg 633
Db      179  GCAAGGGAAGCAAAATGCATATGATTTATCATGAGATTTCTTAATTAATTAATACATTG 238
Oy      634  gttaggggaaagaa-agaagcctcaatgagtgagggagcaagacagagatcgcaattgtcg 692
Db      239  GTAGGGGAAACAGACAGCTCAATGATGATGAGGCGCAAGAACAGAGATGCAATTGCTCG 298
Oy      693  tgccttagtcgaaaccccaagattctgatttgaatgagctacgtctgccttgatc 752
Db      299  TGCCTTAGTTCGAAACCCCAAGATTCTGTATTTAGATGAGCGTACGCTGCCTGGATTTC 358
Oy      753  agaagcaagtcagctgttcaagctgacactgagagagagcgagc 795
Db      359  AGAAMCGAGTCAGCTGTTCAAGCTGCACTGAGAGAGATACC 401

RESULT 4
AA243820      405 bp      mRNA      EST      06-AUG-1997
LOCUS         z167906.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:568506 5',
DEFINITION   similar to SW:MDRL_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ; ,
              mRNA sequence..
ACCESSION    AA243820
VERSION      AA243820.1 GI:1874631
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 405)
              Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
              Kucaba, T., Lacey, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
              Schellenberg, K., Stepien, M., Tan, F., Theisling, B., White, Y., Wyllie
              T., Waterston, R. and Wilson, R.
              WashU-Merck EST Project 1997
              Unpublished (1997)
              Contact: Wilson RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@wustl.edu
              This clone is available royalty-free through LNLN ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Insert Length: 1572 Std Error: 0.00
              Seq primer: -28m13 rev2 ET from Amersham.
              Location/Qualifiers
                1..405
                /organism="Homo sapiens"
                /db_xref="GDB:5562473"
                /db_xref="taxon:9606"
                /clone_image="568506"
                /clone_lib="Soares_NHMPu_S1"
                /tissue_type="Pooled human melanocyte, fetal heart, and
                pregnant uterus"
                /lab_host="DH10B"
                /note="Organ: mixed (see below); Vector: pTR73D-Pac
                (Pharmacia) with a modified polylinker; Site_1: Not I;
                Site_2: Eco RI; Equal amounts of plasmid DNA from three

```



cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10334 row: n column: 21  
 High quality sequence stop: 650.

FEATURES  
 source Location/Qualifiers

1. 1019  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4488404"  
 /clone\_1ib="NCI\_CGAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="3 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: PCMV-SPORT6; Site.1: Salt;  
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 276 a 262 c 283 g 196 t 2 others  
 ORIGIN

Query Match 9.6%; Score 273.2; DB 11; Length 1019;

Best Local Similarity 63.2%; Pred. No. 3,6e-56;  
 Matches 437; Conservative 0; Mismatches 253; Indels 2; Gaps 1;

2116 gaaggaattatgagttcgaagaagctctctctctcattccatgctgcccagatgtttc 2175  
 4 GAAGGAATTATGAGTTTGTGAGAGTCGTCGTCACACTATCCACCCGACCCAGATCCCA 63  
 2176 atccctcgtgctatccctcaglatgagcgaggaagacagatagcttctgtggagaa 2235  
 64 GTGCTTCAGAGGCGCTGAGCCTTGAGTGAGGAAGGCGCAGACCGCTGCGGCGAGC 123  
 2236 agcgcgtgtgggaagaagcctctgttcaactctgtcagagatattatgcccgagaa 2295  
 124 AGTGCCTGCGGGAAGAGCACTGTGCTGAGGAGGAGGCGCTTTCAGACCCCATGGCT 183  
 2296 ggaacagctgttctgagtggtgagtgcaagaagaattgaatgacgtgctcgtcc 2355  
 184 GGATCAGTCTTCTAGATGCGCAAGAAATTAAGCAACTAATGTCAGTGGCTCCAGCA 243  
 2356 caaatagcaatcgttctcctcaagagcctgtctctcaactgacagctgtctgaaacatc 2415  
 244 CAGCTGGCATTGTGTCCCAAGAGCCCATTTCTTGTGACTGCGAGCATCGCAGAACTT 303  
 2416 gcttatgtggaagaagccgtgtgtgtccattagatgagatacaagaagccgaatgca 2475  
 304 GCTTACGGAGACACAGCGGGTCTGTCTTATGAGGAGATTGTGAGGCGCAAGAGAG 363  
 2476 gcaaatatcattcttcttgaaggtctccctgagaatacaacaagaattgagcgt 2535  
 364 GCCAATCCACACAGTTCATGCACTCGTACTGATTAATACACCCAGATAGAGAGC 423  
 2536 aagaagagcagccttctgtgagcgcaagaacaagaatgactgtgcaagggctcttc 2595  
 424 AAGAGCAGCTCAGCTGTGCGGGGCGAGCAGCGCATCGCATCGCAGCGCCCTCGTC 483  
 2596 caaaaaaccataatttatgttgaatgagcgaactcgaagccctcgatataagcagtgag 2655  
 484 AGACAGCCTCACTATTTACTTCTTGAGCAGAGCAACATCCTGTGATACGAAATGGA 543  
 2656 aaggtgttccagatgacctgataaagcagagcaggaaggaacatgagtggtcact 2715  
 544 AAGGTGTTCAGGAAGCGCTGGAACAAGCGAGGAAGCGCCGACCTGTGATGTCGCT 603  
 2716 caccagctctctgcaattgagaagcagattgtagtagtgtt--ctgcacaattggaaga 2773  
 604 CACCGCTGTCCACATCCAGAAAGCGGAGCTTGATGAGTGATTCAGAAAGCGGAAG 663

2774 taaaggaacaaggaactcatcaagagctctg 2805  
 664 TCAAGGAGACGAGCGCACCCCAACAGCAAGTG 695

RESULT 7  
 LOCUS BF584668 894 bp mRNA EST 12-DEC-2000  
 DEFINITION 602098406P1 NCI\_CGAP\_Co24 Mus musculus cDNA clone IMAGE:4218385 5',  
 mRNA sequence.  
 ACCESSION BF584668  
 VERSION BF584668  
 KEYWORDS BF584668.1 GI:11658386  
 SOURCE EST  
 ORGANISM house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 894)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM9798 row: 1 column: 02  
 High quality sequence stop: 651.

FEATURES  
 source Location/Qualifiers

1. 894  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4218385"  
 /clone\_1ib="NCI\_CGAP\_Co24"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: colon; Vector: PCMV-SPORT6; Site.1: NotI;  
 Site.2: Salt; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.6 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 BASE COUNT 263 a 203 c 242 g 185 t 1 others  
 ORIGIN

Query Match 9.5%; Score 270.8; DB 11; Length 894;  
 Best Local Similarity 67.0%; Pred. No. 1,4e-55;  
 Matches 428; Conservative 0; Mismatches 208; Indels 3; Gaps 3;

166 attatgataaagaacccagatagataaacttccacagctgatatataacctgatacc 225  
 106 ATATTGATTAATATAGCCAGTATAGACAGCTTCTCAAGAAGTGGGCAACAACAGCAAC 165  
 226 atagaagaacgtgtggaattaaatgttcttcaatataatcaatcaagaccatctatc 285  
 166 ATACAAGGAATCTGGAATTTAAGATATTCACCTTACCTTACCACTCGAATAAGAGTT 225  
 286 aagatctgaaagtgatgaactcagaattaaatgctggagagcagctgctgtgcgt 345  
 226 CAGATCTTAAGAGGCTCTCAATCTGAAGGTGAAGAGCGGACAGCGGTGGCTGTGTGGC 285  
 346 ctcaatgagcagtggaaggtacggtgagctcgaagcttgcagaggtatatagtccgat 405  
 286 AACAGTGGCTGTGAAAAGCAACAATGTCCACTGATGGAAGGCTCTACAGACCCCTTA 345  
 406 gatgcttatcatatgtgtgagtgagaaatgaatcaagaagcctttaaagtgcggaatcga 465  
 346 GATGCGATGTCAGTATCGACGACAGCAACATCAGAACCATCAATGTGAGGTATCGAGG 405  
 466 gaccatattggagtggtgtgtaagaagcctgttctgttggagaccacatcagtaacat 525



```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CUAAX09"
/clone_lib="Cu"
/tissue_type="adrenal cortico adenoma for Cushing's
syndrome"
/dev_stage="Adult"
/notes="Vector: pBluescript sk(-)"
BASE COUNT      210 a      180 c      211 g      183 t      1 others
ORIGIN

```

```

Query Match      9.2%; Score 262.2; DB 10; Length 785;
Best Local Similarity 64.0%; Pred. No. 1.7e-53;
Matches 458; Conservative 0; Mismatches 254; Indels 4; Gaps 4;

Oy 1939 ggcacgttcacagttttactgcaatgacatgagcatgagccatcgaaacgctc 1998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GGAATTGGAGAGATTTTTCAGCTGTTGCTTTGTCATGGCCATGGCCGGGCAAGTCAGT 132

Oy 1999 gtttgctcctgatatccaaagcgaatcgaggctgcatgctgttgctgttg 2058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 TCATTTCCTCGACTATGCAAAAGCAAAATATCATGAGCCCATCATCATGATCAT 192

Oy 2059 gaaaagaaccaatatalagacagccgagtcacagaaggaaagccagacacatgtgaa 2118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 GAAAAAACCCCTTTTGATTTGACAGCTACAGCAGGAAAGCCCTATATGCCGAAACATTTGAA 252

Oy 2119 gggaaattagattgacgaagctctctctctctctctctctctctctctctctctc 2178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 GGAATATGCATATTGTGGAAGTTGATTTCAATCTCCACCCGACCGAGATCCAGTG 312

Oy 2179 ctccgtgctatccctcaatgacatgacgagaaagacagatgacatgctgtggagcagc 2238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 CTTCAGGAGCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 372

Oy 2239 ggcctgagaaagacactcctgctcaactctgacagacacttctgacccgtgcaaga 2298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 GCGTGTGGAGAGAGACAGCTGCTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 432

Oy 2299 caagtcgttctgacgtgctgacgaagaaatgaaatgacgtgctcctgctccaa 2358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 AAAGTGCTGCTTGATGAGCAAAAGAAATTAAGGACATGATGTCAGTGCTCGAGACAC 492

Oy 2359 atagaaatcgttccccaagagcctgtgctctcaactcagacatgctgtgagaaatgccc 2418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 493 CTGGGATCGTGTCCAGAGAGCCCATCTGTTGACTGACGATGCTGAGAACTTTGCGC 552

Oy 2419 tatgtgacacaagcc-gtgtgtgacatlagatgagatcaagaagccgcaaatgacgc 2477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 TATGAGAGACACAGCCGGGGTGTGTCTACAGAGAAAGATTGTGAGGGCAGCAAGAGAGCG 612

Oy 2478 aaataatcattcttatttgaaggtccctcctgagaaatacaacacacagttgacgtgaa 2537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 613 CAACATACATGCTTCATTCAGACTGACGTGCT-ATAAATATACACTTAACATNGAGACAA 671

Oy 2538 agggagacagcttctgctgagcagaaacaaagacatgactatgcaaggctctctcca 2597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 672 AGGACCTAGCTCTCTGTTGGGGCCAGAAACACACA-TTGGCAATAGCTTGTGGCCCTGTGT 730

Oy 2598 aaaacccaatttattgttggatgaggaacactcagccctcctgataatgacagtg 2653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 CAGACTTTATAA-TTTGTTTTTGATGAAGCCAGTCACTCTTGATACAGAAAGTG 785

```

```

RESULT 10
BE749379 560 bp mRNA EST 25-APR-2001
LOCUS DEFINITION 200194 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE749379
VERSION BE749379.1 GI:10163371
KEYWORDS EST.
SOURCE cow.

```

```

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 560)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Pietra,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCAGACG
Plate: 108 row: F column: 11
Seq primer: ATTAGTGACACTATAG.
FEATURES
source 1..560
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPORT6; site_1: XbaI; site_2: XhoI;
library made from pooled tissue from day 20 and day 40
embryos."
BASE COUNT 154 a 153 c 146 g 107 t
ORIGIN
Query Match 8.8%; Score 252.4; DB 10; Length 560;
Best Local Similarity 67.5%; Pred. No. 4.1e-51;
Matches 355; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Oy 2325 aaaagaattgaatgacagtgctcgtcccaatagaaatgctcctcaagagcctgt 2384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 AAAGAAGCTCAACGTCACAGTGCTCAGAGCCCACTTGAATGCTGTGACAGAGCCCGT 60

Oy 2385 gctctcaactgacagcattgtctgagaacatgctcctatggtgacacagccgtgtgtgc 2444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 CCTGTTGACCTGACCACTATGCGGACAACTTGCATGCGGACCAACAGCGGCTGTAA 120

Oy 2445 atagatagatcaagaagccgcaaatgacgaatcatctcttattgaaggtct 2504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CATGCCCTGAATTTGAGGCGACGCCAAGACGCAACATCTCTTTTCATTTAGACCTT 180

Oy 2505 ccttgaagaatacaacacacagttgacatgagaagagacagcttctgctgagcagaa 2564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GCCCACAATATGAACAAGAGTGGAAGATTAAGGGAGCTCAAGCTTCCGGGGAGACAGA 240

Oy 2565 acaaaagactagctatgcaaggctctctccaaaacccaatttattgtgtgagta 2624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 ACAGAGGATTTGCTATTGGCCGAGCCCTCATCCGACACCCGCCCATCTACTGCTGATGA 300

Oy 2625 ggcacactcagccctcgtataatgacagtgagaggtgtgtgacatgccccttgataaagc 2684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 AGCCAGCTGACGACTGATGAAAGTAGAGAGATTGTCCAGAGAGCCCTTGACAAAGC 360

Oy 2685 caggacgggaagacatgctagtgctactacagagctctctgcaatcagaacgcaga 2744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 CCGAGAGGGCCGACCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 420

```

QY	2745	ttataatggtcttcggcacaatgaagaataagaacaagaactatcaagagcttc	2804
Db	421	CTTGATAGTGGTGATTCGAAATGGACAGCTACGAGGACGACGACACCCAGCACTGCT	480
QY	2805	gagaatcgcagacatatatttcaagtttgatgtaatgcacagtcagtg	2850
Db	481	GGCAGACGAAGAAGGCATCTATTCCACATGCTGTCAGTGTCCAGGCTGGG	526
RESULT	11		
BF969667	BF969667	981 bp	mRNA
LOCUS	602272046p1	NIH_MGC_84	Homo sapiens
DEFINITION	cDNA clone IMAGE:4360090 5', mRNA sequence.		
ACCESSION	BF969667		
VERSION	BF969667.1	GI:12336882	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 981)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I. M. A. G. E. Consortium (LNLN). DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I. M. A. G. E. Consortium/LNLN at: http://image.lnl.nih.gov Plate: LNLN1000 row: 1 column: 11 High quality sequence stop: 670. Location/Qualifiers 1..981 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4360090" /clone_id="NIH_MGC_84" /tissue_type="adrenal cortex carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site: 1; Note: Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1,229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."		
FEATURES	source		
BASE COUNT	267 a 194 c 250 g 270 t		
ORIGIN			
Query Match	8.7%; Score 249.4; DB 11; Length 981;		
Best Local Similarity	61.4%; Pred. No. 2.4e-50;		
Matches	400; Conservative 0; Mismatches 251; Indels 0; Gaps 0;		
QY	1394	tggcttcaagccatgttatatcaggtatctgcttggttgatgaataaagaaacagca	1453
Db	2	TGGTTTCCGATGCAATGCTCAGACAGAGAGTGTGGTTGTTATGACCTTAACCAACCA	61
QY	1454	caggagagcttgcacaacatatattgacatagatatagcacaaatcaaggagcaaggt	1513
Db	62	CTGGAGCATTTGACTACACAGGCTCGCCAAATGATGCTGCTCAATTTAAAGGGCTATAGTT	121
QY	1514	ccaggattgctcttaacaacaaatgcaactaacaatggagacttcagttatcattccot	1573
Db	122	CCAGGCTTCTGTAAATTACCAAGAAATATATGCAAACTTGGGACAGGAATATATATCT	181
QY	1574	ttatatatgtagtgagagatgacatctcgtatctcgaagtatgctccagtaactgcccgtga	1633
Db	182	TCATCTATGTTGGCAACATACATCTTATCTTGTAGCATTTTACCAATCATTTAAATAG	241

Y	1634	caggatgatgtaaacccgcagcgaatgactgggtttgccaacaagaatagaagaactta	1693
Db	242	CAGAGATTGTTGGAATGCAAAATGTTGTCTGCGACAGCACTGAAATATGAAAGAACTAG	301
Qy	1694	agcatgtcggaaagatagacactgaactgaacctttggaagaataacgtaactatagtgatc	1753
Db	302	AAGTGCTGGGAAGATCGTACTGAAAGCAATGAAAGAACTCCGAACCGTTGTTCTTTGA	361
Qy	1754	caaggggaaaaagccttcgagcaaatgtatagaagagatgcttgaagctaacacaggaata	1813
Db	362	CTCAGGAGCAGAACTTGAAACATATGTATGCTCAGAGTGTGGAGGTACCATACAAACT	421
Qy	1814	cctgaagaagaagcagagattattgggaagcgtgtatgcatctggacgatgoccttataat	1873
Db	422	CTTGAGAGAAAGCAACACATCTTTGGATTATACATTTCCCTTACCCAGCAATGATGATT	481
Qy	1874	tgtcctatgcagcagaggtttcgaatttggagcctatttaattcaagctggagcaatgacc	1933
Db	482	TTTCTGTATGCTGATGTTTCCGGTTTGGAGCCCTACTGTGTCGACATTAACATCAGACT	541
Qy	1934	cagagggcaatgttcatagttttttacctgcgaattgcataatgagctatgccaatcggaana	1993
Db	542	TTTGAGAGATGTTCTGTATGTATTTTCAGCCCTGTTGTCCTTGTGTCACATGGCCGAG	601
Qy	1994	cgctcgctttggcctccgaatatctccaagaacccagcgggcgtgcgacatc	2044
Db	602	TCAATTGATTTGCTCCTGACTATGCAAGCAAAATATACAGACGCAATC	652
RESULT	12		
BFJ13560			
LOCUS	BFJ13560	795 bp	MRNA
DEFINITION	601900129FI NIH_MGC_19	Homo sapiens	cdna clone IMAGE:4129222 5',
VERSION	BFJ13560		
KEYWORDS	BFJ13560.1	GI:11261583	
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 795)		
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgads-femail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov		
	Plate: LLCM1023 row: h column: 23		
	High quality sequence stop: 674.		
FEATURES	Location/Qualifiers		
source	1..795		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4129222"		
	/clone_1ib="NIH_MGC_19"		
	/tissue_type="neuroblastoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: brain; Vector: pORF7; site:1; XhoI: site_2		
	ECORI; cDNA made by oligo-dT priming. Directionally		
	cloned into EORI/XhoI sites using the following 5'		
	adaptor: GGCAAGG(G). Library constructed by Ling Hong		
	in the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies).		
	Note: This is a NIH_MGC library."		
BASE COUNT	213 a	189 g	157 t



ORIGIN

Query Match	8.7%;	Score 249;	DB 11;	Length 795;
Best Local Similarity	62.1%;	Pred. NO. 2.9e-50;		
Matches 479;	Conservative	0;	Mismatches 280;	Indels 12; Gaps 5,

QY	2064	gaacacaaatataagacacgcgcagctcaagaagaaggaagaaacagacatctgaaggaa	2123
Db	1	GAACCCCTTTGATGACAGCTACAGCACGGAAGGCTTAATGCCGAACACATGGAAGAAAT	60
QY	2124	tttagaagtttcgaagaagtcctcttctctcatccatcgtcgcgccagatglttctcatcccg	2183
Db	61	GTCACATTTGGTGAAGTGT---ATTCAACTATATCCACCCGACCGGACATGCCAGTGTCTTCA	117
QY	2184	tgccttatccctcagatattgagcgcgaggaagacagtagcatttctgtgggagcagcgtc	2243
Db	118	GGGAGCTGAGCCTTGAGAGGTGAGAAAGGGCCAGACGCTGGCTGTGTTGGGACAGAGTGGCTG	177
QY	2244	tgggaaagaagcattcgttcaactcttgcagaagactttagtgcccccgtgcgaagacaa	2303
Db	178	TGGGAAGAGCACAGTGGTCCAGTCTCTGGAAGGGTTTCTTAGCACCCCTTTGGCGAGGAAGT	237
QY	2304	gcgttttgaatggtgtgattgcacaaagaatltgaatgtacagtgtgcgcgttctcccaatagc	2363
Db	238	GCATGCTTGATGGCAAGAATAAAGGAGCTGATGTTCAGTGTGGTCCAGACACACTGG	297
QY	2364	aatcgttctcctcaagacgtcgtgctcttccaactctgcagcattgtctggaacatcgcctatg	2423
Db	298	CATCGTGTCCACAGAGACCCATCCTCTGTTTACAGTCAGCATGCTTGAGAACATTCGCTTAGG	357
QY	2424	tgcacaaagcgcgtgtgtgctccatttaataagtttcaagaagcgcgaataatgcagcaaat	2483
Db	358	AGACAAACAGCGGGGTGGTGTCAACAGGAAGAAGTCTGAGCGACGAAGAAGAGCCCAAT	417
QY	2484	ccattcttattgaagagttccctctagaataatcaacacacaaagtgtgaactgaagaagc	2543
Db	418	ACATGCTTATCGAAGTCACTGCTTAATTAATATACCATTAAGTAGAGACAAAGAAC	477
QY	2544	acagcttctcgcgcgcgaagaacaaagactagctattgcaagggctctctcctcaaaaac	2603
Db	478	TCACCTCTGTGGTGGCCAGAACAAAGCATTTGGCATAGTCTGTGCCCTTTGATGACACCC	537
QY	2604	caaaatttatgttggatgagcgcactcttcagcccctcgtaatgacatgtgaagatg-y	2663
Db	538	TCATATTTTGGT--TTGGATTAAMCCACAGTCAGCTTGATTAAGAAAGTAAAGGTGCG	595
QY	2663	ttcagcagtcgcctctgataaagccagagcagcggaagagacatgcctagtgtgcataccagc	2722
Db	596	TCCAGAAGCCCTGAGCAAAGCCA-----GAGAAAGGAGACCTGCATGTATGTGTACCCG	651
QY	2723	tctctgcaattcagacgcagatttgatagtgttctgcacaaatgnaaataaagaac	2783
Db	652	TGTCACCAATTCAGAAATGAACATT--ATAGTGGTGTTTCAGAAATGGACAGATCCAGGGG	709
QY	2783	aagaactatctcaagagctctctgaagaatcgcgaacataattttaagttag	2833
Db	710	CATGACGACATCAAGAGTGTCTGTGACGAAGAAGCTCATTTTCTACGCTAG	760

TITLE

JOURNAL  
COMMENT

Determination of clone end sequences from *Entamoeba histolytica*  
HMI:IMSS sheared DNA library (2001)  
Unpublished (2001)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: enta@tigr.org  
Clones are derived from the *Entamoeba histolytica* HMI:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 18  
High quality sequence stop: 851.

FEATURES  
SOURCE

BASE COUNT  
ORIGIN

Query Match	8.7%;	Score 248.8;	DB 13;	Length 886;
Best Local Similarity	-58.4%;	Pred. No. 3.4e-50;		
Matches 458;	Conservative 0;	Mismatches 317;	Indels 9;	Gaps 1;

Oy	176	agaaccacgatatgatacaactcttccacagctgtagatataaacctggaatccatcgaatgaaggaa	235
Db	47	AGAAAAACCCACTATTGATTTGATTGTATGTGAGAAAGGTGAAACATTTAAATGATGTTAAAGGTG	106
Oy	236	ctctggaatttaaaaaagtcttccttcaatataccataagaaccactcataagaattcga	295
Db	107	AAATTGAAATTTAAAGACACTTTGTTTCAGATATCCAAACAGACCGACAAATTCGTCTTGA	166
Oy	296	aagctctcgaatcactagaattaaactcgtgagagagacagtcgccttgctgcgtcctaagca	355
Db	167	AAGGATTTTCATTCGAAGTAGAACAAAGSAAAAACTGTTGCTTGTAGTAGACATCAGAT	226
Oy	356	gtgggaagaagtcaggttagccagctctctcgagaggtatataatgacggaatgagcttta	415
Db	227	GTCGTAATCAACACTACGCTTCAGTTGATTTGAACAAATTTTATGATCCAAACATGGACATG	286
Oy	416	tcaatggtgataagagaatgaactcagagcttaaatgtggcgcatatccgaaacatatgtg	475
Db	287	TATTATTGATGGACATPATATATCAAAGTTTGAATTTCAATTTCTTAAGAACTCAAAATTCG	346
Oy	476	gaatggttagtcagaagagcgtcttgcttcggaccacatcagtaaacataatacaagtatg	535
Db	347	GAATGCTAGGCAACAAAGACCACTATTATTTCGTGAAGTGTTATGATATATTTATGAAGAC	406
Oy	536	g-----acgagatgcttgactgataagaagatvgagagagcgcaagaaggaacaa	566
Db	407	GAGTACCTAAAGAGAGTCGTAAGTATATGAACAATTTATGCTGCTGCTAAATGCGCA	466
Oy	567	atgcgtatgatttatactgagttccctaaattataactctgtctgagggaaaag	646



Db 467 ATGCACATGACTTATTTCAGCAATGCCAGAGATATACCAATAGGTAGTATAGAG 526  
 Qy 647 gagctcaaatgagtgagggagagaaacagagatcgcaattgcctgctagtgtaa 706  
 Db 527 GTGCACAAATTTCAAGAGAGCAAAAACAAGAAATTTGCTATTGACAGTCAATTGATGAA 586  
 Qy 707 accccaagattctgattttagatgagctgctgccttgatcagaagcaagtcag 766  
 Db 587 ATCCAAAGTGTTATTACTGATGAGTACATGACACTTGATTCAGAAAGTGAAGA 646  
 Qy 767 ctgtcgaagcgcagctgagagagcgagcaagtcgagctacaatcgctgtagacacc 826  
 Db 647 TTGTACAGAGTACACTTACAAAGAGCAAAAAGAGAACACAACTTGTATTGACATA 706  
 Qy 827 gacttctactatcgaagtcgaattgattgtagccctaaagagagaaatgctgagcg 886  
 Db 707 GATTTCACACTTTCAAATATGCAATATGCTTATTATGAGAGGAGAAATTTGCAG 766  
 Qy 887 agaaagagacacatgctgactaataatgcaaaacgagctctatattacttgatgt 946  
 Db 767 AAAGAGGAACACATCAAGAGTTATGATTTGAAGAGATTTTATATACACTTGCTATGC 826  
 Qy 947 caca 950  
 Db 827 AACA 830

RESULT 14  
 LOCUS BG587938 780 bp mRNA EST 11-APR-2001  
 DEFINITION EST1489713 MHAM Medicago truncatula/Glomus versiforme mixed EST  
 library cDNA clone PMHAM-51L19 5' end, mRNA sequence.  
 ACCESSION BG587938  
 VERSION BG587938.1 GI:13603002  
 SOURCE Medicago truncatula/Glomus versiforme mixed EST library.  
 ORGANISM Medicago truncatula/Glomus versiforme mixed EST library  
 Eukaryota; mixed EST libraries.  
 REFERENCE 1 (bases 1 to 780)  
 HARRISON, M.J., LIU, J., TOWN, C.D., VAN AKEN, S., UTTERBACK, T., CHO, J.  
 AND FRASER, C.M.  
 ESTs from roots of Medicago truncatula after colonization with  
 Glomus versiforme, 2001  
 Unpublished (2001)  
 CONTACT: HARRISON M.J.  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73401  
 Tel: 580-223-5810  
 Fax: 580-221-7380  
 Email: mjharrison@noble.org  
 Noble EST name: N387011e TIGR sequence name: MTDDK707K More  
 information is available at: <http://www.medicago.org>  
 Seq primer: SKmod (CTA gaa cta gtc gat cc).  
 Location/Qualifiers  
 1..780

FEATURES  
 source  
 /organism="Medicago truncatula/Glomus versiforme mixed EST  
 library"  
 /cultivar="Medicago truncatula genotype A17"  
 /db\_xref="taxon:119092"  
 /clone="PMHAM-51L19"  
 /clone\_lib="MHAM"  
 /tissue\_type="roots colonized with Glomus versiforme"  
 /dex\_stage="roots harvested at 10, 17, 22, 31 and 38 days  
 post-inoculation with Glomus versiforme. The library was  
 made from a mixture of RNA from each of these stages."  
 /lab\_host="E. coli strain XLOLR"  
 /note="Vector: pBluescript SK-. Site 1: EcoRI; Site 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA from  
 roots harvested at 10, 17, 22, 31 and 38 days  
 post-inoculation with Glomus versiforme. The cDNA was  
 directionally ligated into the Unizap XR vector from  
 Stragene and packaged using Gigapack III Gold packaging

BASE COUNT 234 a 156 c 184 g 206 t  
 ORIGIN  
 extracts. Plasmids containing cDNA inserts were excised  
 from the recombinant lambda-Zap phage using Ex-assist  
 helper phage and propagated in XLOLR cells."

Query Match 8.5%; Score 242.6; DB 11; Length 780;  
 Best Local Similarity 58.5%; Pred. No. 1,1e-48;  
 Matches 444; Conservative 0; Mismatches 309; Indels 6; Gaps 1;

Qy 2009 ctgaatattccaagaacaaatcgcgggctgcgcacatgctgttgccttlttgaaagaac 2068  
 Db 28 CTGCATTTCACAAAGCGTGCAGCTTGCCTGCTAAGATTTCGGATTTATTCATCACAGC 87  
 Qy 2069 caaatatagacagccgcagcttaagaaggaaacccagacatgtaaggaattag 2128  
 Db 88 CTGGTATGATAGAAACGTGAATCTGGATTGGAATTAAGACAGTTTACTGGACTTGTG 147  
 Qy 2129 agttcgaagaagtccttctctatccatgctgcgccagatgttccatccctgctgct 2188  
 Db 148 AACTCAAAAATGTGACTTCTCTTATCCATCAAGACCTGAAGTTCTATCTCATATGATT 207  
 Qy 2189 tatccctcagatitgaagcgaagaaagacagatagcaatttggaggagcggcgtgga 2248  
 Db 208 TCTCCTTGAAGTTCCTCCGGAAGAACCATACCTTTAGTTGATGACGCGCTCTGGCA 267  
 Qy 2249 aagagactctgttcaacttctgcagagactttatgaccccgtagaaggaaatgctgt 2308  
 Db 268 AGAGCACTGTTGCTCTTATGAGAGATTCTTGATGACCACTTCAGACAAAGTAATGT 327  
 Qy 2309 ttgagtgtgtgagtcgaagaatgaatgaaatgacagtgctgccttcccaatagaacatcg 2368  
 Db 328 TGAGATGGCAGTACACTTAAACTTTGAAGCTTAATGTTAGGCAACAAATAGACTAG 387  
 Qy 2369 ttccctcgaagcctgtgcttctcaactgcagcattgctgagaacatgcctcctatgtgaca 2428  
 Db 388 TGAGCCAAAGAACCTGCTTTGTTGGCCACGAGATTGAGAAAATATCTTGGGA---- 443  
 Qy 2429 acagccgctgtgtgctcatatagatagatcagaagccgcaaatgagcaaatatcat 2488  
 Db 444 --AGCCTGATGCAAAACAGGTTGAGATTGAAGAGCTGCTAGGGTTCCTAATGCTCAT 501  
 Qy 2489 ctcttatgaaggtctccctctagaataatacaacacaaatgtgagctgaaagagacacagc 2548  
 Db 502 CATTCATATCAAACTTCTCTGAAGCTTTGAACACTCAGAGTGAGAAAGAGACTACAC 561  
 Qy 2549 ttctcgcgcgcagagaacaaagactagctatctgcaagggctcttccaaaacccaaa 2608  
 Db 562 TTTCTGCGAGGACAAAACAAAGAAATAGCAATAGCAAGGCGCATGCTAATAAATCCAGCA 621  
 Qy 2609 ttctattgttgatgagccacttccagccctcgatgaatgaagaagtgagagtggttcgc 2668  
 Db 622 TTTCTCTCTCGATGAGGCAACAACTGATGACTTCATGCAATACGAAAGGTGGGCAAG 681  
 Qy 2669 atgccttgataaagccgcgagcgaagagacatgctagtgatcactcaacagctctctg 2728  
 Db 682 AAGCACTTGACCGGTTCAATGATGGCAGACACACTCTGTAATGTGCTATGCGCTCTCCA 741  
 Qy 2729 caattcagaacgacagattgtagtgcgttcgcacaaatg 2767  
 Db 742 CAATTCGCAAGCTGACCTGTAGCGCGGATTCACAAAG 780

RESULT 15  
 LOCUS A2540627 886 bp DNA GSS 14-NOV-2000  
 DEFINITION ENT00187F Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 genomic, DNA sequence.  
 ACCESSION A2540627  
 VERSION A2540627.1 GI:11147603  
 KEYWORDS GSS.  
 SOURCE Entamoeba histolytica.

Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.

JOURNAL

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0200

Fax: 301 838 0208  
Email: [enta@tigr.org](mailto:enta@tigr.org)

Clones are derived from the *Entamoeba histolytica* HM:IMSS sheared DNA library  
Seq primer: M3-Forward  
Class: shotgun

## Source

1. 886  
/organism="Entamoeba histolytica"  
/strain="HMI:IMSS"  
/db.xref="taxon:5759"  
/clone.lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pMOSt; Site: 1; Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Using a DNA isolated from broth cultures of *E. histolytica*  
genomic method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) *Entamoeba histolytica*: a  
method for isolate identification. *Exp. Parasitol.*  
77:450.) The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In *Genome  
Sequencing: A Practical Approach*, eds. M. Vaubin and B.  
Barrell, Oxford University Press, 1999)."

BASE COUNT	364 a	107 c	181 g	234 t
ORIGIN				

Query Match	8.5%;	Score 241.4;	DB 13;	Length 886;
Best Local Similarity	59.5%;	Pred. NO. 2.2e-48;		
Matches 449; Conservative	0;	Mismatches 296;	Indels 10;	Gaps 2

QY	169	atgtataagaagaccacgatatagataaactttccacagcttgatataaaccgtatcatc	228
Db	22	ATTGATATGCAATACCAATATTTGATTTGATATCTACACCTGGATTAGTTCCAACTGACTGT	81
QY	229	gaagaacctgtagaatttaaaatgcttccttcaatcattccataagaaccatcatcaag	288
Db	82	ATTGGAATATTTAAATATAGAAAGATGTTCAATTTAGATATCCAAACAAACCAATTAACAA	141
QY	289	attctgaaaagtcgaatctccagaattaaatctcggagaagcagtcgccttgctgcgtctc	348
Db	142	ATTCTTGCTGCACTTGACCTTGAATATTAAGAAAGAAACAAACAGTTGATTATGATGAGCA	201
QY	349	aatgycagttggaagagtaacgtaagtcacagcttcctgcagaagttataatgaaccgaatg	408
Db	202	TCAAGATGTGTGAATGCACACTACAAATTCATTTAGTCCAAAGAGTCATGATCCAGTTGCT	261
QY	409	gaccttatcatgtagatctgtagaatagatacaagagccttaaatctgcgcattacgaagac	468
Db	262	GGAAAGATCAATTAGATGGAAAAAGATTTAAGAGAGTTGAACCTTGAATGTTAAAGAAAT	321
QY	469	catatctgagtggttaatcaagaagccgtttctgtcggaagccacatcagtaacaatc	528
Db	322	CAATATAGATTAAGTTGGACAAAGAAACATTTTGTTCATGCTACTATTAAGAAATAT	381
QY	529	aagtagagcagagatgatg-----tgactgatagaagagatbgagagagcaagaag	579

Db	382	ATGCTTGAGAGCTAAAGATGGAGAAACACCAACTGACACAGACAGATGATGAAATGCTGCTAA	441
QY	580	gaagcaaatgcgtatgatttlatcatcagaggttccataataatlaataatlgtagg	639
Db	442	ATGCAATGACATGAAATTTATTTCTCATCTTCACAGAGATGATGACAAATGGTTAGCA	5010
QY	640	gaaaaaggagctcaaatgagtggagggcagaacaagagatcgcaattgctgtgctta	699
Db	502	GAAGAGAGAGCTGATTTATATCAGGAGGACAAAACAAAAGAAATTGGTATTTGACGCGCAATTC	561
QY	700	gttcgaacccccgaagtattgattttagatggagctacgtctgcccggattcgaanagc	759
Db	562	ATTGAAAAACCTACAAATTTTATTACTTGATGAGAACCTCACTACAGACTTGATACCAAGT	621
QY	760	aagtcagctgttcaagctgcacgtcagagaaagcgagcaaaagtcgagctaacatcgttga	819
Db	622	GAAAGAAATTGTACAAACCAAGCACTTGAAAAAGCTGACAGGAAGAACACAAATTTATTTGA	681
QY	820	gcacacgcgaacttctactatctgaagtgcagatttgattgtg-acccctaaagatggat	878
Db	682	GCACATGATTTAACCAACGTTGAGAAATGCAAAATTTGGAATTTGTGTATTCATCAAGAGCA	741
QY	879	gtctcgagaaaggaagcagctgtcgaactaatg	913
Db	742	AATTAATTGACACAGAACACCTCAAGATTTATATGG	776

Search completed: April 22, 2002, 18:31:47  
Job time: 12200 sec





```

0Y 226 atgaagaactgctggaattaaatgtttcttcataatcacaagaaccatctac 285
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1300 ATTAGGGAATTTGGAAATTCAGAAATGTTCACTTAGTTTCCATCTCGAAAGAAAGTT 1358
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 286 aagaattcgaagaagctctgaaatccagaataaagtcctgagagacagctgccttggcgt 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1360 AAGATCTTTGAAGGGCCCTGAACTGAAAGTGCAAGATGGGCACACGGGTGGCCCTGGTTGGA 1419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 346 ctcaatggcagtggaagaagtacgtgtagcttcagcttcctgcagaagttatgatctggat 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1420 AACATGTGCTGTGGGAAGACACCAACAGTCTCAGTATGCAGAGGCTGTATATACCCACACA 1479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 406 gatgcttcalcatcagtgagatgaagaatgacatcacagctttaaattgctgcattatcga 465
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1480 GAGGGGATGGTCACTGTTGTATGGACACAGAGATATTAGACCATTAATATGTAAGTTTCTACGG 1539
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 466 gaccataatggagtggcttgaatcaagagccttgcttctgctggagcaacaatcagtaacaat 525
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1540 GAAATCATTTGGTGTGGGAGACTGACGGAACCTGTATTTGTTGCCACACACATAGTACTGAAAC 1599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 526 atcaagaatgacagagatgtgtgactgatatgaagaagatgagagagacagcaagggaaaga 585
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1600 ATTGGCTATGTGCGCTGAAATATGTCACCATGAGATGAGATTTGAGAAAGCTGTCAAGAGACC 1658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 586 aatgcgatgatttaltcatcagtgagttccctaataattaatatcattgtagtgagaaaaa 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1660 AATGCTTATGACTTTATCATGAAACATGCCCTCATTAATTTGACACCCCTGTTGAGAGAGA 1719
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 646 ggaagctaaattgaatgtagagggcagaagaagagatgcgaattgctgcgttacttagttcga 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1720 GGGGCCAGTTGAGATGGTGGGACGAAGACAGAGATGCAATGCAAGTCCCTGGTTGCG 1779
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 706 aacccaagaattctgattttagatgtgagctacgtctgcacctgattatcagaagaacagtlca 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1780 AACCCCAAGATCCTCTGTGATGAGAGCCAGCATGCGCTTGAGACAGAAAGCAAGACA 1839
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
0Y 766 gctgttcaagctgcacgcgaggaag 790
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1840 GTGCTTCAAGTGGCTCTGTGATTAAG 1864
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-08-784-649A-5
: Sequence 5, Application US/08784649A
: Patent No. 5830697
: GENERAL INFORMATION:
: APPLICANT: SIKIC, Branimir I
: APPLICANT: Chen, Gang
: TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
: TITLE OF INVENTION: CYCLOSPORIN MODULATION
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 2200 Sand Hill Road
: CITY: Menlo Park
: STATE: CA
: COUNTRY: USA
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/784,649A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Sherwood, Pamela J
: REGISTRATION NUMBER: Reg. No. 5830697 36,677
: REFERENCE/DOCKET NUMBER: 06037/007001
: TELECOMMUNICATION INFORMATION:

```

```

: TELEPHONE: 415-322-5070
:
: TELEFAX: 415-854-0875
:
: INFORMATION FOR SEQ ID NO:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH: 4264 base pairs
:
:     TYPE: nucleic acid
:
:     STRANDEDNESS: single
:
:     TOPOLOGY: linear
:
:   MOLECULE TYPE: cdna
:
:
: JS-08-1784-649A-5

```

Query Match	27.0%;	Score 317.8;	DB 2;	Length 4264;
Best Local Similarity	-69.3%;	Pred. NO. 7.3e-88;		
Matches 433;	Conservative 0;	Mismatches 192;	Indels 0;	Gaps 0;

[illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525



Db	1526	ATATTTGATATATAAGCCAAAGTATTGACAGGTATTCGAAGAGTGGGACAAACCAAGATAT	1583
Oy	226	atagaagaacgtctgtgaatttaaataatgttcttcttcaattatccatcaagaccatctc	285
Db	1586	ATTATAGGGAATAATTGGAAATTCACAAAATGTTCATCTTCAGTTACCCATTCGAAATAAGATT	1645
Oy	286	aagattctgaaagtcgtgaatctcagaattaaagtcctgagagacagtcgccttgctgt	345
Db	1646	AAGATCTTTGAAGGCCCTCGAAGCTCGAAGGTCAGAGTGGGCAAGCGTGCCCTGGTTTGA	1705
Oy	346	ctcaatgycagttggaagagtaagtcgaagtcacgcttctgcgaagttatgatccgagat	405
Db	1706	AACAGTGGCTGTGGGAGAGCACACAAACAGTCCAGCTGATCAGAGGCTCATATGACCCACA	1765
Oy	406	gatygccttatcatgctgctgagatggaatgatcatcaagagccttaaatgctgcgcatatcga	465
Db	1766	GAGGGGATGTGCACATGTTGATGACAGCAGATATTAGGACCATTAATGATAAGTTGTATACGG	1825
Oy	466	gaccatttgagagtggttaagtaagaagccgtgttcttgctggagaccacatcaagtaacat	525
Db	1826	GAATTCATTGGTGTGGTGAGTCAAGAACCTGTATTGTTTGGCCACCAAGTAACTTAACAAAC	1885
Oy	526	atcaagatggaagcagatga tgcactgatatgaagatgaagatgaagagacgaagggaaaga	585
Db	1886	ATTGGCTATGGCCGTGAATAATGTCCACATGATGAGATTTGAGAAAGCTGTCAAGAGAGCC	1945
Oy	586	aatgcgatgatgattatcatcatgagagtttccctaataaattaatatcatattgtagggaaaaa	645
Db	1946	AATGCTATGACTTTATTCATGAAGAACTGCCATCAATAATTTTGACACCCCTGGTTGGAGAGGA	2005
Oy	646	ggaagctcaaatgaattgagagggcagaagaacagaagatcgaactgtcgtgagccttagtga	705
Db	2006	GGGGCCAGTTGAAGTGGTGGGCGAAGACAGAGATGCGCATTTGCACAGTCCCTGGTTGCC	2065
Oy	706	aaccoccaagatctgattttagatlgaggtcagctgtcccttgatctcaagaagcaagta	765
Db	2066	AACCCCAAGATCCTCTGCTGAGATGAGAGCCAGCTCAGCCTTGAGACAGAAAGCAAGACA	2125
Oy	766	gctgttcaagctgcagctgaggaag 790	
Db	2126	GTGTTCAAGTGTGCTCTGATTAAGG 2150	

RESULT 5  
 5206352-3  
 Patent No. 5206352  
 APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,  
 Michael M.  
 TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA  
 SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
 NUMBER OF SEQUENCES: 4  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/622,836  
 FILING DATE: 24-SEP-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 892,575  
 FILING DATE: 01-AUG-1986  
 APPLICATION NUMBER: 845,610  
 FILING DATE: 28-MAR-1986  
 SEQ ID NO:3:  
 LENGTH: 4669  
 5206352-3

Query Match	27.0%	Score 317.8	DB 6	Length 4669
Best Local Similarity	69.3%	Pred. 7.7	E 88	
Matches 433	Conservative 0	Mismatches 192	Indels 0	Gaps 0

  

QY	166	atatttgaagaagaaacccggtatagataactttccacagcgtgatataaacctgaatcc	225
DB	1526	ataattgataataaagcaagatctgacagctatctgaagaagtgggacacacccaataat	1585
QY	226	atagaagaactctggaatttaaaaagtcttttcattatcctcaagaacatcattac	285

Db	1586	atlaagggaatttgcgaattccagaaatgttccattccattgctattcccatccatccgaaagaagtt	1645
Qy	286	aaagattctgaaagtcctgaaatccagaatataagtcctggaagaacagtcctgtgtgctgt	345
Db	1646	aaagattctgaaagtgccgaaacctgaaaggtgcagagtggtgcagaaagcgtgtgacctgtgtga	1705
Qy	346	ctcaatgacagctgggaagagtagcagtagtccagcttcacagaggtatatagtatccgatt	405
Db	1706	aaacgtgtgcctgtgggaagagcaacaacagtcacgtctgtatgcaagcgtctatgacccaca	1765
Qy	406	gatgcttatacatbgtgtgagtgagaatagacatcagaagctttaaattgtgcgcatatbga	465
Db	1766	gaagggagatgtcagtgctgtcttgatgacagatataatagacatataagatttatacag	1825
Qy	466	gacctatttgagtggttaatacgaagcctgtttgtctgggaacacccatcaatgaacat	525
Db	1826	gaatactctgtgtgtgtgaaatccgaaacctgtatgtcttgccacacagatagctgaaac	1885
Qy	526	atcaagatgtagcagaga ttgattgtactgatagaaagatggaagagagcagcaagggaaqa	585
Db	1886	attgcgtatgcccgttgaaatgtatccatgatgagatgagatgagaagcgtgtcaaggaagcc	1945
Qy	586	aatgcgtatgatttatacatatgtaggatttccctaataaattaatatgattgttagggaaaaa	645
Db	1946	aatgccttatgacttatacatcaagaacatgacctcatataaatttgacaacctgtgtgagagaga	2005
Qy	646	ggaagctcaaatgaa ttgagaggcagaaacaagaagatcgcgaatgtgcgtgtgccttaattga	705
Db	2006	ggggcccgcttgaa ttgtgtgtggcagaaacaaagaaatctgcattgtgcaagtgccctgtgtgc	2065
Qy	706	aaacccaagattctgaaattttaga ttgagctacgtctgtcccttgaaatcagaaagaatga	765
Db	2066	aaacccaagatccctcctgtgcgtatgagcgaagtcagctcttgacacagaagaagcgaagca	2125
Qy	766	gctgttcaagctgcatcgtgggaag	790
Db	2126	gttgtcaagtgctctctgataag	2150

RESULT 6  
US-08-793-610-5  
Sequence 5, Application US/08793610  
Patent No. 5858744  
GENERAL INFORMATION:  
APPLICANT: BAUM, Christopher  
APPLICANT: STOCKING-HARBERS, Carol  
APPLICANT: OSTERFAG, Wolfgang  
TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
TITLE OF INVENTION: FOR GENE TRANSFER  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/793,610  
FILING DATE: 07-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 31 973.8  
FILING DATE: 08-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 195 03 952.1  
FILING DATE: 07-FEB-1995  
PRIOR APPLICATION DATA:

```

1  APPLICATION NUMBER:  PCP/EP95/031215
2  FILING DATE:  10-AUG-1995
3
4  ATTORNEY/AGENT INFORMATION:
5
6  NAME:  Betman, Richard J.
7  REGISTRATION NUMBER:  39,105
8  REFERENCE/DOCKET NUMBER:  P1614-7000
9  TELECOMMUNICATION INFORMATION:
10
11  TELEPHONE:  (202)638-5000
12
13  TELEFAX:  (202)638-4810
14
15  INFORMATION FOR SEQ ID NO:  5:
16
17  SEQUENCE CHARACTERISTICS:
18
19  LENGTH:  6505 base pairs
20  TYPE:  nucleic acid
21  STRANDEDNESS:  double
22  TOPOLOGY:  circular
23
24  MOLECULE TYPE:  DNA
25
26  OS-08-793-610-5

```

Query Match	27.0%	Score	317.8	DB 2	Length	6505			
Best Local Similarity	69.3%	Pred. No.	9.3e-88						
Matches	433	Conservative	0	Mismatches	192	Indels	0	Gaps	0

Oy	166	atctatgaataagaaaccccaattatagtaactcttcccaagcgtggatataaacttcaatcc	225
Db	2918	ATATATTATATATTAAGCCAACTATTGACACAGTATTGCAAGATGGGCACAAACCAATATAT	2977
Oy	226	atagaagaacactgtyggaatttaaaaaatgtcttccttaactatcatcaagaaccatactac	285
Db	2978	ATTAAGGCAAAATTTGGAAATTCAGAAATGTTCACCTCAGTTACCACATCTCGAAAAAGAGTT	3037
Oy	286	aagatcttcgaaagtgcttgaaatctcagaataatgaatctctgagagacagtgccctgtgcgt	345
Db	3038	AAAGTCTTGAAGGGCCCTGAACCTGTAAGGGTGCACAGAGGGCCAAACGGTGGCCCTGGTTTGA	3097
Oy	346	ctcaatgacagctggaagaagtaacgtaagtcacgaactcttcgaaggattatgaatccgcat	405
Db	3098	AACAGTGGCTGTGGGAAGACACAAACAGTCCACACTATGCAAGGCTCTATATACCCACACA	3157
Oy	406	gatygccttatcatcgtgtgtagtgaagaatgcaatcagaagccttaaatgtgcgcataatcga	465
Db	3158	GAGGGAGTGGCTAGTGTGTGATGAGACAGATATTAGGACCATTAATGTAAAGTTTCTACGG	3217
Oy	466	gaccatacttggaatggtttatgataagaagccgtttttcttggaacaacatacgaatacat	525
Db	3218	GAAAATCATTTGGTGTGGGAGACTCGAGAACCTGTATTTTGGCCACACACGATAGCTTAANAAC	3277
Oy	526	atcaagaatgagacagatgtatgtgacttgatgtaagaagatgtgagagacaagaaggaagca	585
Db	3278	ATTGCGTATGCGCGTGAAATATGTCACACATGAGATTTGAGAAAGCTGTCAAGGAAGCC	3337
Oy	586	aatgcgtatgatttatcatcagtagctttccataaataatcatatcatgttgtaagggaagaa	645
Db	3338	AATGCCATATGACTTTATCAATGAACATGCGCTCATATAATTTTGACACACCTGTGGTGAAGAGA	3397
Oy	646	ggaagctaaatagtgtaggagtgagaaacagaagatcgaactgcgtgcgttccttagttcga	705
Db	3398	GGGGCCACAGTTTGAGTGGTGGGCGAAGAACAGAGAGATGCGCATTTGCACAGTGCCTCGGTTGCC	3457
Oy	706	aaccccaagaattctgattttatgatagtgagctacgcttgcctctgatttcagaagaagcaatca	765
Db	3458	AACCCCAAGATCTCTGCTGATGAGGCGACGATCGACTTGAGACACAGAAAGCGAAGGA	3517
Oy	766	gctgtttcaagctcagctcgtggaag 790	
Db	3518	GTGGTTTCAGGTGAGTCTGTGATTAAG 3542	

```

RESULT 7
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:

```

APPLICANT: BAUM, Christopher  
 APPLICANT: STOCKING-HARRERS, Carol  
 APPLICANT: OSTERTRAC, Wolfram  
 TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
 TITLE OF INVENTION: FOR GENE TRANSFER  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
 STREET: 655 Fifteenth Street N.W. Suite 330  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005-5701  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/793,610  
 FILING DATE: 07-MAR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE P 44 31 973.8  
 FILING DATE: 08-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 195 03 952.1  
 FILING DATE: 07-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP95/03175  
 FILING DATE: 10-AUG-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Berman, Richard J.  
 REGISTRATION NUMBER: 39,105  
 REFERENCE/DOCKET NUMBER: PL614-7007  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)638-5000  
 TELEFAX: (202)638-4810  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9318 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA

[illegible]



Db 3177 GAATCATGTTGTTGTCAGTACAGAACTGTATTGTTGGCCACCAGATAGTGAAGAAC 3236  
Qy 526 atcaagatgagcagatgatgtagcagatgaagaagatgagagcagcaggaagca 585  
Db 3337 ATTCGATATGCGCCGTGAAGAAATGTCACATGATGATGATGAAAGCTGCAAGGAAGCC 3296  
Qy 586 aatgcgatgattatcatcagatgattcctaataatlaatacatgtaggaggaanaa 645  
Db 3297 AATCCATATGACTTTATCATATAAAGTGCATATTAATTTGACCCCTGTTGAGAGAGA 3356  
Qy 646 ggaagctaaatgagtgagagcagaagaagagatcgcaattgctgtagcttagtoga 705  
Db 3357 GGGGCGGAGTGTAGTGGGCGGAGAGAGATCGCATTCGACCGCTGTTGCGC 3416  
Qy 706 aaccacgaatctcagatttagatgagctacgctgcccctgagattgaagaagatga 765  
Db 3417 AACCCAGAGATCTCTCTGCTGATGAGGCCAGTCAGCTTGACACAGAAAGGAGACA 3476  
Qy 766 gctgtcaagctgacatggaagaag 790  
Db 3477 GTGGTTCAGGTGCTCTGATGAAG 3501

RESULT 8  
US-08-583-276-18  
; Sequence 18, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:  
; APPLICANT: Mcdonagh, Kevin T.  
; APPLICANT: Nienhuis, Arthur  
; APPLICANT: Tolstoshev, Paul  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Bain, Giffillan,  
; ADDRESSEE: Cecchi & Stewart  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: DNA V2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,276  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/332,444  
; FILING DATE: 31-OCT-1994  
; APPLICATION NUMBER: 07/887,712  
; FILING DATE: 22-MAY-1992  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4669 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: singular  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: Genomic DNA  
; US-08-583-276-18

Query Match 26.8%; Score 314.6; DB 2; Length 4669;  
Best Local Similarity 69.0%; Pred. No. 7.5e-87;  
Matches 431; Conservative 0; Mismatches 194; Indels 0; Gaps 0;  
Qy 166 atattgataagaaccagatagataactttccacagctgagatataaacctgacatcc 225

Db 1526 ATAATTGATATTAAGCCAGATATGACAGCTATTGCAAGAGTGGGCAAAACGATTAAT 1585  
Qy 226 atagaagaactgcygaatttaaaaaatgcttcttcaattatcatcaagccatcatc 285  
Db 1586 ATTAAGGGAATTTTGAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAATTTG 1645  
Qy 286 aagattcgaagagctgcaatcagaatgaatgaatgaatgaatgaatgaatgaatgaat 345  
Db 1646 AAGATCTTGAAGGCGCTGACCTGAGGTGAGAGTGGGCGAGAGGTGCCCTCGTTGGA 1705  
Qy 346 ctcaatggcagtggaagagtagcgttagtccagctcttcagagaggtatataatgacgat 405  
Db 1706 AACAGTGGCTGGGAAGAGACAAACAGTCCAGCTGATGACAGAGCTTATGACCCACA 1765  
Qy 406 gatgcttatacatcagtgagatgaagaatgaatgaatgaatgaatgaatgaatgaat 465  
Db 1766 GAGGGATGCTAGTGTGATGACAGAGATTTGAGACCATTAATTAATTAATTAATTAAT 1825  
Qy 466 gaccatattgagctgtagttagcagaagcctgtttgttcgagacacacatcagtaacaat 525  
Db 1826 GAATCATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1885  
Qy 526 atcaagatgagcagatgtagttagttagttagttagttagttagttagttagttagt 585  
Db 1886 ATTCGCTATGCGCCGTGAAGAAATGTCACCATGATGATGATGATGATGATGATGATG 1945  
Qy 586 aatgcgatgattatcatcagatgattcctaataatlaatacatgtaggaggaanaa 645  
Db 1946 AATGCTATGACTTTATCATATAAAGTGCATATTAATTTGACCCCTGTTGAGAGAGA 2005  
Qy 646 ggaagctaaatgagtgagagcagaagaagagatcgcaattgctgtagcttagtoga 705  
Db 2006 GGGGCGGAGTGTAGTGGGCGGAGAGAGATCGCATTCGACCGCTGTTGCGC 2065  
Qy 706 aaccacgaatctcagatttagatgagctacgctgcccctgagattgaagaagatga 765  
Db 2066 AACCCAGAGATCTCTCTGCTGATGAGGCCAGTCAGCTTGACACAGAAAGGAGACA 2125  
Qy 766 gctgtcaagctgacatggaagaag 790  
Db 2126 GTGGTTCAGGTGCTCTGATGAAG 2150

RESULT 9  
US-09-120-513-1  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Ellens, Harma  
; APPLICANT: Field, John  
; APPLICANT: Yue, Lin  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MOR1B2 AND  
; TITLE OF INVENTION: SCREENING METHODS THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/120,513  
; FILING DATE: 22-JUL-1998

```
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: King, William T
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: GP50008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5015
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4233 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-120-513-1
```

```
Query Match 26.3%; Score 309.6; DB 3; Length 4233;
Best Local Similarity 68.3%; Pred. No. 2.5e-85;
Matches 429; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
```

```
QY 166 attatgtaagaagaaccagatagataactttccacagctgagatataaactgaatcc 225
DB 1121 ATATTTGATATGAGCGACGACATGTGACAGCTTCTCAACCAAGGACACAAACGACAGT 1180
QY 226 atagaaggaactgtggaatttaaaatgttcttcaattatccaaagaccatcacc 285
DB 1181 ATATTTGGGAAATTTGGCAATTTAAATGTTTACTTCACTACCCACATCAAGAACTGAAGT 1240
QY 286 aagattctgaaggtctgacttcgaatgaatgctcgagagagacgtcgttcgtgcgt 345
DB 1241 AAGATCTTGAAGGCGCTCACTCAAGTGAAGAGCGGCGACGCTACCGCTGTTGGC 1300
QY 346 ctcaatgagcgtggaagaagtaacgtgactgcagctctgcagaggtatataatccgagat 405
DB 1301 AACAGTGGCTGTGGGAAAGACAACTGTCCAGCTCTGCAAGAGCGCTTACGACCCCATTA 1360
QY 406 gatgcttatcatcagtgtgataagaaatgacatcagaagcctttaatgtcgcgcatatcga 465
DB 1361 GAGGGGAGAGTCAATTCACGAGACAGACATCAGACCATCAATGTGAGTCTGTCGGG 1420
QY 466 gaccatattgaggtgtgactgaagaagcctgtttgttcggagacccatcagtaacaat 525
DB 1421 GAAATCTATGGGCTGTGAGTCAAGGACCGCTGTGTCACCAAGATGCGCGAAAC 1480
QY 526 atcaagtaatgagcagatgactgactgataagaagatgagagagacaaagaaggaagca 585
DB 1481 ATTTCGTATGGCGAGAAACGTCACATGATGATGATGATGATGATGATGATGATGATGAT 1540
QY 586 aatgcgtatgacttcatcagtgaattcctaataaatttaatacatgtgtaagggaagaa 645
DB 1541 AATGCTTATGACTTCATCTGAAACCTGCCCAATTTAACCCCTGTTGGTGTGAGAGA 1600
QY 646 ggaagcctaataagtgagagagcagaagaacagagatgcgaattgtcgttcgttcagtg 705
DB 1601 GGGGGCGACCTAGTGGGGGACAGAAACAGAGGATGCGCATTTGCCGCGGCGCTGTCGCC 1660
QY 706 aaccccaaatcttgatttagatgaggtacgtctgcgcgtgattcaaaaagaagcaagta 765
DB 1661 AACCCCAAAATCTTTTGTGTGATGAGGCGCACGCTCAGCTTGGACACAGAAAGCGAAGCC 1720
QY 766 gctgttcaagctgacgtgagaagata 793
DB 1721 GTGTTCAAGCGCGCTCTGTGATTAAGCTTA 1748
```

```
RESULT 10
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
```

```
APPLICANT: Kimberly Anne Brun
APPLICANT: Richard James Chenery
APPLICANT: Harma Ellens
APPLICANT: John Anthony Feld
APPLICANT: Lin Yue
TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
FILE REFERENCE: GP-50008-D1
CURRENT APPLICATION NUMBER: US/09/450,105
EARLIER FILING DATE: 1999-11-29
EARLIER APPLICATION NUMBER: 09/120,513
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 4233
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-450-105-1
```

```
Query Match 26.3%; Score 309.6; DB 4; Length 4233;
Best Local Similarity 68.3%; Pred. No. 2.5e-85;
Matches 429; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
```

```
QY 166 attatgtaagaagaaccagatagataactttccacagctgagatataaactgaatcc 225
DB 1121 ATATTTGATATGAGCGACGACATGTGACAGCTTCTCAACCAAGGACACAAACGACAGT 1180
QY 226 atagaaggaactgtggaatttaaaatgttcttcaattatccaaagaccatcacc 285
DB 1181 ATATTTGGGAAATTTGGCAATTTAAATGTTTACTTCACTACCCACATCAAGAACTGAAGT 1240
QY 286 aagattctgaaggtctgacttcgaatgaatgctcgagagagacgtcgttcgtgcgt 345
DB 1241 AAGATCTTGAAGGCGCTCACTCAAGTGAAGAGCGGCGACGCTACCGCTGTTGGC 1300
QY 346 ctcaatgagcgtggaagaagtaacgtgactgcagctctgcagaggtatataatccgagat 405
DB 1301 AACAGTGGCTGTGGGAAAGACAACTGTCCAGCTCTGCAAGAGCGCTTACGACCCCATTA 1360
QY 406 gatgcttatcatcagtgtgataagaaatgacatcagaagcctttaatgtcgcgcatatcga 465
DB 1361 GAGGGGAGAGTCAATTCACGAGACAGACATCAGACCATCAATGTGAGTCTGTCGGG 1420
QY 466 gaccatattgaggtgtgactgaagaagcctgtttgttcggagacccatcagtaacaat 525
DB 1421 GAAATCTATGGGCTGTGAGTCAAGGACCGCTGTGTCACCAAGATGCGCGAAAC 1480
QY 526 atcaagtaatgagcagatgactgactgataagaagatgagagagacaaagaaggaagca 585
DB 1481 ATTTCGTATGGCGAGAAACGTCACATGATGATGATGATGATGATGATGATGATGATGAT 1540
QY 586 aatgcgtatgacttcatcagtgaattcctaataaatttaatacatgtgtaagggaagaa 645
DB 1541 AATGCTTATGACTTCATCTGAAACCTGCCCAATTTAACCCCTGTTGGTGTGAGAGA 1600
QY 646 ggaagcctaataagtgagagagcagaagaacagagatgcgaattgtcgttcgttcagtg 705
DB 1601 GGGGGCGACCTAGTGGGGGACAGAAACAGAGGATGCGCATTTGCCGCGGCGCTGTCGCC 1660
QY 706 aaccccaaatcttgatttagatgaggtacgtctgcgcgtgattcaaaaagaagcaagta 765
DB 1661 AACCCCAAAATCTTTTGTGTGATGAGGCGCACGCTCAGCTTGGACACAGAAAGCGAAGCC 1720
QY 766 gctgttcaagctgacgtgagaagata 793
DB 1721 GTGTTCAAGCGCGCTCTGTGATTAAGCTTA 1748
```

```
RESULT 11
US-08-461-823-1
; Sequence 1, Application US/08461823
```

```

; Patent No. 5593840
; GENERAL INFORMATION:
; APPLICANT: Bhatnagar, Satish K.
; APPLICANT: George Jr., Albert L.
; APPLICANT: Nazarenko, Irina
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OncorPharm, Inc.
; STREET: 200 Perry Parkway
; CITY: Gaithersburg
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20877
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,823
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/168,621
; FILING DATE: 16-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,433
; FILING DATE: 27-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Karta, Glenn E.
; REGISTRATION NUMBER: 30,649
; REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301 527-2058
; TELEFAX: 301 208-6997
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-461-823-1

```

```

Query Match      18.8%; Score 220.6; DB 1; Length 2726;
Best Local Similarity 58.6%; Pred. No. 5.3e-58;
Matches 404; Conservative 0; Mismatches 279; Indels 6; Gaps 1;

```

```

QY 108 ctgtaggttctttagtgtaacatcagcaggtatttgatgagcagcagctccat 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1477 CAGTTCATTCTCTCTGCTATGCAAGCCAAATATCAGCAGCCACATCATATGAT 1536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 tatgataagaagccagatagataactttccacagctgagataaactgaaatcat 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1537 CATTAAGAAAACCCCTTGATGACGACGACGACGACGACGACGACGACGACGAC 1596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 228 agaaggaactgtgaaattaaatgattcttcaattatcatcaagaacatcataaa 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1597 GGAAGGAATGTCACATATTGCTGAGATTGTATCAACTATCCACCGACGACATCCC 1656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 288 gatttgaaggtctgaaatcagaatgaagtctgagaagacagtcgcttggtgct 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1657 AGTGCTTACGAGGACTGAGCTGGAGGTGAAGAGGCGCAGACGCTGCTGCTGGCGAG 1716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 caatggaagtggaagagtagtagtccagctctgcagaggtatagatgacgagatga 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1717 CAGTGGCTGTGGGAAGACACAGTGCTCAGCTCTGGAGCGGCTTCTACAGACCCCTTGGC 1776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 408 tggcttatcatgtgtgagatgagaatcagagctttaaatgtgcgcatatcagaga 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 1777 AGGGAAGTGTCTGCTGATGGCAAGAAATTAAGCGACTGAATGTTAGTGCAGGC 1836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 468 ccatatgagtagtgtagtaagaagccgtgttctgcggagccaccatcagtaacaatat 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1837 ACACCTGGGCACTGCTGCTCCAGAGACCCATCTCTTTGACGACGATGCTGGAACAT 1896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 528 caatgtaga-----cgagatgtagtctgactgatagaagatgagaagagaagga 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1897 TGCCATGAGACAAACAGCCGGGTGTGTACAGGAAGAGATCTGAGAGCGACGAAGA 1956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 582 agcaatgcgtatgattatcaatgaggttccctaataattatcatatggtgagga 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1957 GGCCAACTATACATGCTTCATCGAGTCACTGCTTAATTAATAGCACTAAAGTAGAGA 2016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 642 aaaagagctcaaatgagtgagaggcagaacagagagatcgcaattgctgctgcttagt 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2017 CAAGGAAGCTACGCTCTGTGTGGCCAGAAACAAACGCTTCCATAGCTCGTCCCTTGT 2076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 702 tcgaaccccaagattctgattttagatgaggtctagctgctgagatcagaagaacaa 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2077 TAGACAGCCTCATATATTGCTTTGGATGAAGCCACGTCACCTCTGGATACAGAAAGTGA 2136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 762 gtcaagctgtcaagctcactgagagaag 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2137 AAGGTGTCCAGAAAGCCCTGAGCAAG 2165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atld of
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:

```

NAME/KEY: CDS  
LOCATION: 1..4002  
US-08-996-545-1

Query Match 14.6%; Score 171; DB 2; Length 4002;  
Best Local Similarity 55.8%; Pred. No. 1.3e-42;  
Matches 349; Conservative 0; Mismatches 270; Indels 6; Gaps 1;

QY 178 aaacccagatataactttccacagcttgatataaacctgaatccatagaagaact 237  
DB 3205 AAGCCCAATATGATACGTCTGTAAGAGGCGGAGAGAGTCAAAACGGGAGAGGAA 3264  
QY 238 gtggaattaaatgtttcttcaattatccatgaagaccatctcaagattcga 297  
DB 3265 ATGGAATTTAGGAGAGTGCATTCACATTAACCCAGCCGCGCAACAGCTTCTGCGC 3324  
QY 298 gtgtgaatctcagaatlaagtcgagagacagtcgcttgcgttcgaatgcag 357  
DB 3325 GCGTTGACCTGACCGTGAAGCTGGAACAATATGTTGCGTTGCGAGCCAGCGGTGT 3384  
QY 358 ggaagagtaagtgatccagcttcgagaggtatagatccgagatgagcttacc 417  
DB 3385 GGCAGAGATACACCAATTCATTCGTGAGCGCTTTACGATGCGATTGCCGGTCCATC 3444  
QY 418 atgtgtgagagatgacatccagacttaaatgtgcgcatatccagacatattg 477  
DB 3445 CTGTGTGATGAGGAGAGACATTAAGTAACTAATTAATCAACTCTTACCGACCTTCTGTCA 3504  
QY 478 gtgtgttagtaagaagcctgtttgttcgagacacatccagatatacaatgaatg- 536  
DB 3505 CTGCTGACGAGGAGCGACACTGTACAGGAGCACCATCAAGAAAACATCTTACTTGTT 3564  
QY 537 -----acgagatgattgagatgagagagagagagagagagagagagagagag 591  
DB 3565 ATTTGTGAGAGTACGCTACCGAGAGATTTCTTATTAAGCTTTCAGAGGACGCTAATATC 3624  
QY 592 tatgattttatcagtgagtttcttaataatlaatacatctgtgtaagagagagagag 651  
DB 3625 TACGACTTATCATGTGCTCCCGAGAGGCTTTAATTAATCACTTGTGACAGAGGAGGC 3684  
QY 652 caaatgagtgag 711  
DB 3685 ATGTTGTGTCGGCGGCAAAAGCAAGCGTGCCATTCGCCGAGCCCTTTCGGAGATCC 3744  
QY 712 aagattctatctttagatgagagagagagagagagagagagagagagagagagag 771  
DB 3745 AAAATCTTCTTCTCGATGAGGAGGAGTCAAGCCCTGACACTCGAGTCAAGAAAAGTCTGTC 3804  
QY 772 caagctgactgag 796  
DB 3805 CAGGCGGCTTTGATGAGCGCTGCGC 3829

RESULT 13  
US-08-996-545-3  
Sequence 3, Application US/08996545  
Patent No. 5928898  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Maard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atld of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,545  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-996-545-3

Query Match 14.6%; Score 171; DB 2; Length 4002;  
Best Local Similarity 41.9%; Pred. No. 1.3e-42;  
Matches 262; Conservative 87; Mismatches 270; Indels 6; Gaps 1;

QY 178 aaacccagatataactttccacagcttgatataaacctgaatccatagaagaact 237  
DB 3205 AAGCCCAATATGATACGTCTGTAAGAGGCGGAGAGAGTCAAAACGGGAGAGGAA 3264  
QY 238 gtggaattaaatgtttcttcaattatccatgaagaccatctcaagattcga 297  
DB 3265 AUGGAAUUGAGAGAGGAGGACUACAGUACCCAGCCGCGCAACAGCCUUGUCGCGC 3324  
QY 298 gtgtgaatctcagaatlaagtcgagagacagtcgcttgcgttcgaatgcag 357  
DB 3325 GCGUUGAGACCGGAGGAGCGGAGCAUUGUUGCGUUGCGAGCCAGCGGUGU 3384  
QY 358 ggaagagatcagtgatccagcttcgagaggtatagatccgagatgagcttacc 417  
DB 3385 GGCAGAGATACACCAATTCATTCGTGAGCGCTTTACGATGCGATTGCCGGTCCATC 3444  
QY 418 atgtgtgagagatgacatccagacttaaatgtgcgcatatccagacatattg 477  
DB 3445 CUGGUGAUGGAGAGGAGCAGUACUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3504  
QY 478 gtgtgttagtaagaagcctgtttgttcgagacacacatccagtaaatcaatgaatg- 536  
DB 3505 CUGGUGAGCCAGGAGCGGAGCAGUACUAGGAGGAGCAUUAAGAGAAACUUAUUGGU 3564  
QY 537 -----acgagatgattgagatgagagagagagagagagagagagagagagagag 591  
DB 3565 AUGGUGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3624  
QY 592 tatgattttatcagtgagtttcttaataatlaatacatctgtgtaagagagagagag 651  
DB 3625 UAGCAGUACUACUAGUUGGCGCCGAGAGGCGUUAUUAUUAUUAUUAUUAUUAUUAUUA 3684  
QY 652 caaatgagtgag 711  
DB 3685 AUGUUGUCUGGCGGCAAAAGCAAGCGUGGCGCAUUGGCCGAGCCUUAUUGGAGUCC 3744  
QY 712 aagattctatctttagatgagagagagagagagagagagagagagagagagagag 771  
DB 3745 AAAAUUCCUUAUUGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3804  
QY 772 caagctgactgag 796

Db 3805 CAGCGCGCUUGAGUCCGCCGCC 3829

# RESULT 14

US-09-328-320-1  
Sequence 1, Application US/09328320  
Patent No. 6228615

## GENERAL INFORMATION:

APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of  
TITLE OF INVENTION: Aspergillus nidulans  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/328,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/996,545  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763

## INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4002  
US-09-328-320-1

Query Match 14.6%; Score 171; DB 4; Length 4002;  
Best Local Similarity 55.8%; Pred. No. 1.3e-42;  
Matches 349; Conservative 0; Mismatches 270; Indels 6; Gaps 1;

QY 178 aaaccagatagataacttcacagctgatataaaccgaatccatagaagaact 237  
Db 3205 AAGCCCAATGATGATGCTGTGAGAGCGGAGAGCTGAAAGGTGAA 3264  
QY 238 gtgaatttaaaatgtttctcaatlaacatcaagaccatctcaagattctga 297  
Db 3265 ATCGAATTATAGAGAGCTGACATACCGAGCCGCCAGAACAGCTGTCTGCC 3324  
QY 298 ggtctgaatctcagaatagtcgtgagagacagtcgtctgtgctgctcaatgagc 357  
Db 3325 GCGTTGAGACCGTACCGTGAAAGCGTGAACAATATGTTGCGCTGTGAGACCCAGCGTGT 3384

QY 358 ggaagaagtagcagtagtccagcttcgcagaggtatataatccgagatgcttacc 417  
Db 3385 GGCAGAGTAGACCACTTGCATTGCTTGAGCGCTTTAGACGAGATTGCCGGTCCATC 3444  
QY 418 atggtgagatgaatgacatcagagctttaaatgtgcgcatlatacagacatttga 477  
Db 3445 CTTGTGATGAGGAGACATAGTAACATAATACACTCCCTACCGCGAGCTTCTGCA 3504  
QY 478 gtgttagtcaagagcctgtttgttcggaccacatcagtaaaataatagaatag 536  
Db 3505 CTGCTGAGCCAGAGCGCACTGTACAGGACCATCAAGAAATCACTTACTTGCT 3564  
QY 537 -----acgaatgatgtgactgatgaagaatgagaagcagcaaggaacaaatc 591  
Db 3565 ATTGTGAGATGAGCTTACCGAGATCTTGATTAAGCTTGCAAGAGCCTATATAC 3624  
QY 592 tatgattatcatgagttccctataaattacatcattgtaggggaaagagact 651  
Db 3625 TACGACTTCATCATGCTCCCGAGGCGCTTAATACAGTTGTGCGAGCAAGGAGGC 3684  
QY 652 caaatgagtgagggcagagaacagagatcgcaattgtcgtgcttagtcgaaccc 711  
Db 3685 AGTTGTCTGCGGCCAAAGACGTGTGCGCATTCGCCGAGCCCTTCTTGGAATCC 3744  
QY 712 aagattctgattatgagtagcagctgcctgcctgagattcagaagcaagcagctgt 771  
Db 3745 AAATTCCTTCTTCGATGATACGAGCAGCCTCGACTCGAGTCAAGAAAGTGTCT 3804  
QY 772 caagctgcactgagaagataccc 796  
Db 3805 CAGCGCGCTTGTGAGCCGCTGCC 3829

# RESULT 15

US-09-328-320-3  
Sequence 3, Application US/09328320  
Patent No. 6228615

## GENERAL INFORMATION:

APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of  
TITLE OF INVENTION: Aspergillus nidulans  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/328,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/996,545  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-328-320-3

Query Match 14.6%; Score 171; DB 4; Length 4002;  
Best Local Similarity 41.9%; Pred. No. 1.3e-42;  
Matches 262; Conservative 87; Mismatches 270; Indels 6; Gaps 1;

QY 178 aaaccagatagataactttccacagctgatatataaactgaatccatagaaggact 237  
DB 3205 AAGCCACAATUUGAUAACUGGUCUGAAGAGGCGAAGACUUGCAAAACGGUGAAGGUGAA 3264  
QY 238 gtggaatttaaaagtcttctcaattatcatcaagaacatctatcaagattctgaa 297  
DB 3265 AUCGAUUUAGGAACGUGACUUCAGAUUACCCGACCCGCCAGAACAGCCUUGCCUGCGC 3324  
QY 298 ggtctgaatctcagaatgaatgaatcagtcgagacagtcgcttgctcggtcgaatgagc 357  
DB 3325 GGCUUUGAACUUGACCGUGAAGCCUGGACAAUUVUUGCCUUGCGGACCGCGUUGU 3384  
QY 358 gggaaagatagcgtatgcagcttctgcagaagttatatgatccgagatgattgttacc 417  
DB 3385 GGCAGAGAGACCAACAUUUGCUUGAGCGCUUUGACAUUGCGAUUGCCGGGUCAC 3444  
QY 418 atggtggaatgaatgaatcagagcttcaatgtgagcatatcagacatattgga 477  
DB 3445 CUUGUUGAUGGAGAGACUAUAAGUAAUAUACUCCUACCGCAGCUUUCUGUCA 3504  
QY 478 gtggttagtcaagagctgtttgttcgagaccatcagtaacaataatcaagtatg- 536  
DB 3505 CUGGUCAGCCAGGAGCGCACUUGUACAGGCGCACCAUACAGAAACAUUUCUUGGU 3564  
QY 537 -----acgagatgatgtgactgtatgaagagatgagagcaaggaagcaaatgcg 591  
DB 3565 AUUGUCGAAGAGAGACGUAUACGAGAAUUCUUGAUUAGGCUUGCAAGAGCGCUAAUUC 3624  
QY 592 tatgatttatcatgagattctcctaataattatcatatgtgtagggaaaaaggagct 651  
DB 3625 UACGACUUCUAUUGCCGUCGCCGAGGCUUUAUACAGUUGUUGCGACAGAGGAGGC 3684  
QY 652 caaatgagtgaaggagcaagaacagagatcgcaattgctgcttgaatttcgaaacccc 711  
DB 3685 AUGUUGUCUGGCGCCAAAGAACGUGUGGCGCAUUGCCGAGCCUUCUUGGGAUCC 3744  
QY 712 aagattctgattttagatagagctagctgcccctgagatccagaagaagtcagctgtt 771  
DB 3745 AAUAUCCUUCUUCGUAAGACGACGACCCUUGACUCCGAGUCAGAAAGGUCGUC 3804  
QY 772 caagctgcactgagaagataccc 796  
DB 3805 CAGCGCGCUUUGAUCCCGCUGGCC 3829

Search completed: April 22, 2002, 18:39:46  
Job time: 12569 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:31:47 ; Search time 9392.76 Seconds  
(without alignments)  
1344.238 Million cell updates/sec

Title: US-09-873-409-11

Perfect score: 1175  
Sequence: 1 cccaatccccaataatctc.....taaaatgctgacttatagga 1175

Scoring table: IDENTITY-NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpl:\*  
6: em\_estlba:\*  
7: em\_estro:\*  
8: em\_estrov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	545	46.4	545	10	AL040762 DKFP434C
2	489.8	41.7	533	11	N24315
3	489.8	41.7	533	11	N24315
4	482.8	41.1	559	11	BF692596 Homo sapien
5	279.6	22.8	405	10	AA243820
6	268.2	22.8	894	11	BF584668
7	220.6	18.8	780	11	BF587938
8	220.6	18.8	943	10	BF587938
9	209.2	17.8	886	13	AL520322
10	207.6	17.7	853	13	AL520322
11	207.6	17.7	880	13	AZ540627
12	207.6	17.7	932	13	AZ679805
					AZ679805
					ENTL0522F
					AZ670821
					ENTN0697F

Result No.	Score	Query Match	Length DB	ID	Description
13	203.8	17.3	661	11	BC302675
14	201.6	17.2	575	10	AM342644
15	200.8	17.1	402	11	H82550
16	192.8	16.4	886	13	BH139685
17	192.6	16.4	891	13	AZ682250
18	191.2	16.3	787	11	BC584063
19	188	16.0	1019	11	BC248052
20	185	15.7	750	11	BC585786
21	184.8	15.7	611	11	BC35924
22	184.6	15.7	687	11	BC646725
23	183.6	15.6	897	13	AZ541090
24	183.2	15.6	834	13	AZ548312
25	182.2	15.5	871	13	AZ682350
26	182.2	15.5	947	13	AZ683753
27	181	15.4	860	13	AZ543293
28	180.8	15.4	656	10	AM686102
29	178.8	15.2	823	13	AZ532602
30	172.4	14.7	921	13	AZ687628
31	167.6	14.3	795	11	BF133560
32	167.4	14.2	919	13	AZ690701
33	165.2	14.1	651	10	BE016688
34	165	14.0	785	10	AV709991
35	164.4	14.0	815	11	B1409057
36	164.4	14.0	2676	12	AK014319
37	163.8	13.9	885	13	AZ686798
38	163.4	13.9	842	13	AZ671925
39	161.6	13.8	944	11	BF796582
40	161.2	13.7	541	10	A1722858
41	160.6	13.7	899	13	AZ548799
42	159.6	13.6	878	13	AZ539524
43	158.8	13.5	926	13	AZ685308
44	157.6	13.4	897	13	AZ550221
45	156	13.3	704	10	BE239652

#### ALIGNMENTS

RESULT 1  
AL040762  
LOCUS  
DEFINITION  
DKFP434C1815\_1 434 (synonym: htes3) Homo sapiens cDNA clone  
ACCESSION  
AL040762.1 GI:5409708  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 545)  
REFERENCE  
Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
Wiemann, S.  
EST (Bloecker, et al.)  
Unpublished (1999)  
CONTACT: Bloecker H

FEATURES  
source  
Am Klopferpitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
s1 sequence also available.  
This clone (DKFP434C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFP434C1815"



```

/clonelib="434 (synonym: hies3)"
/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="vector: pSPori1; Site_1: NotI; Site_2: SalI"

```

```

BASE COUNT      171 a      89 c     138 g     147 t
ORIGIN

```

```

Query Match      46.4%; Score 545; DB 10; Length 545;
Best Local Similarity 100.0%; Pred. No. 2.5e-128;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 182 ccagatagaacttccacagcttgatataaacctgaatcacaagaagaaactgtg 241
    |||||||
DB 1  CCAATATAGATTAATTTCCACAGCTGATATAAAGCTGATGAGAGAACTGTGG 60
QY 242 aatttaaaatgttcttcaatataatcacaagaacatcacaagaatcagaagctc 301
    |||||||
DB 61 AATTTAAATGTTCTTCAATATTCATCAAGACCATCATCATCATGATTTCTGAAGGTC 120
QY 302 tgaatcccaaatgaatgtggagagagagctgcctgtgtcgtcgaatgacagtg 361
    |||||||
DB 121 TGAATCTCAGAAATTAAGCTGTGAGAGAGAGTGCCTTGTGCTCAATGCGAGTGGGA 180
QY 362 agaatcagatgctcagctcctcagagatgatatatgacagatgagcttatacag 421
    |||||||
DB 181 AGAATGAGTATGCTCAGCTTCTGAGAGGTATATGATATCCGATGATGCTTATCATAG 240
QY 422 tgaatgagatgacatcagaagcttcaatgtgagcttatacagagacatattgag 481
    |||||||
DB 241 TGGATGAGAAATGACATGAGACCTTTAAATGTGCGCATTAATGAGACCATTTGTGAGTGG 300
QY 482 ttagcaagagccctgttctgtcggagaccacacacatcaatatacagaatagag 541
    |||||||
DB 301 TTAATCAAGAGCCCTGTTTGTTCGGGACACCATCATCAATATCAATGATGAGACAG 360
QY 542 atgatgacatgacagagatgagagagagagagagagagagagagagagagagag 601
    |||||||
DB 361 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 602 tcaatgagatgttccatataatataatgataatgataatgataatgataatgata 661
    |||||||
DB 421 TCATGAGATGTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 662 gagggcagaacagagagatgcgaatgtcgtgcttgaatgcgaacccaagatctga 721
    |||||||
DB 481 GAGGGCAGAAACAGAGATGCCAATGTCTGCTTACTTGAATCCCAAGATTCTGA 540
QY 722 tttaa 726
    |||||
DB 541 TTTTA 545

```

```

RESULT 2
LOCUS N24315 533 bp mRNA EST 28-DEC-1995
DEFINITION Yx33c11.1 Soares melanocyte 2NBHM Homo sapiens cDNA clone
IMAGE:262580.5' similar to gb:M14758 MULTIDRUG RESISTANCE PROTEIN 1
(HUMAN); mRNA sequence.
ACCESSION N24315
VERSION N24315.1 GI:1138465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 533)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
M., Holtzman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston
R., Williamson, A., Wohldmann, P. and Wilson, R.
The WashU-Merck EST Project

```

```

JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu

```

```

High quality sequence stops: 353
Source: IMAGE Consortium, LIND
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.lind.gov) for further information.
Seq primer: 77
High quality sequence stop: 353.

```

```

FEATURES
    source
        Location/Qualifiers
            1..533
                /organism="Homo sapiens"
                /db_xref="GDB:387222"
                /db_xref="taxon:9606"
                /clone="IMAGE:262580"
                /clonelib="Soares melanocyte 2NBHM"
                /sex="Male"
                /issue_type="melanocyte"
                /lab_host="DH10B (ampicillin resistant)"
                /note="vector: p77r3D (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTATCCATCTGATGAGTGGAGCGGCGCGCGCTTTTCTTTTCTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo. RNA from normal foreskin melanocytes
(F8374) was kindly provided by Dr. Anthony P. Albino."

```

```

BASE COUNT      139 a     134 c     118 g     142 t
ORIGIN

```

```

Query Match      41.7%; Score 489.8; DB 11; Length 533;
Best Local Similarity 98.1%; Pred. No. 2.9e-114;
Matches 517; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

```

```

QY 651 tcaatgagatgagagagagagagagagatgcgaatgtcgtgcttgaatgcgaaccc 710
    |||||||
DB 1  TCAATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
QY 711 caagattcgaattgaatgagatgagatgagatgagatgagatgagatgagatgag 770
    |||||||
DB 61 CAAGATTCTGATTTAGATGAGGCTACGCTGCTGCTGATTCAGAAACCAATGAGCTGT 120
QY 771 tcaatgagatgagagagagagagagagagagagagagagagagagagagagag 830
    |||||||
DB 121 TCMACTGACCTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 831 tggagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 890
    |||||||
DB 181 TGGAGATGCTGAGCTTGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 891 ctgacctcctgctgctcctatgagatgagatgagatgagatgagatgagatgagat 950
    |||||||
DB 241 CTGACCTCCTGCTGCTTGAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 300
QY 951 ctgacccaatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1010
    |||||||
DB 301 CTGACCAAAATTAATCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 1011 gacagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 1069
    |||||||
DB 361 GAAAGATTTTTCGAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 420
QY 1070 -ggcttaatgagacatcttgaatgagatgagatgagatgagatgagatgagatgag 1128
    |||||||
DB 421 GGCTTAATGGGACATTTACTTGTGATTTGCTTGGAGAGTGAAGTGAAGGCTTTTTC 480

```

Oy    1129 tctaagaaaaatcgcaggcttctttttaaatgctgacttatgga    1175  
         ||| |  
Db    481 CCTAAGAAATCGCAGGCTTC-TTTTTAAATGCTGGCTTTATGCA    526

[illegible]

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MAPPING

1 (bases 1 to 533)  
Allikmets, R., Gerrard, B., Hutchinson, A. and Dean, M.  
Characterization of the human ABC superfamily: Isolation and  
mapping of 21 new genes using the expressed sequence tags databases  
Hum. Mol. Genet. 5 (10), 1649-1655 (1996)  
97049974

**AUTHORS** Allikmets, R., Gerrard, B. and Dean, M.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (12-AUG-1996) Human Genetics Section, National Cancer  
Institute, NCI-FCRDC, Frederick, MD 21702, USA

FEATURES	Location/Qualifiers
source	1. .533

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="EST422562"
/note="similar to ATP-binding cassette transporter
BASE COUNT      139 a      134 c      118 g      142 t
ORIGIN

```

Query Match	41.7%	Score 489.8;	DB 12;	Length 533;
Best Local Similarity	98.1%;	Pred. No. 2.9e-114;		
Matches 517;	Conservative	0;	Mismatches 7;	Indels 3;
				Gaps 2

Qy	651	710
Qy	tcaatgagtcggagggcagacaacagagatcgcaattgtctgcgtccttagtcgaancc	710
Db	1 TCAATGAGTCGAGGGCAGCAAAACAGAGGATCGCAATTGCTCGTSCCTTAGTCGAAACC	60

QY 711 caagattctgatttagatgagctacgctcgcctggtattcagaagcaagtcagctgt 770  
|||||  
61 CAAGATTCTGATTTTATGATGAGGCTACGCTGCCCCGGATTTCAGAAAGCAAGTCACCTGT 120

Oy 771 tcaagctgcacgtggaagaagtaccgccagatattcatttgactaattcacccaag 830  
Db 121 TCAAGCTGCACGTGAGAAGCATCCCCCAGTATTTCATTTTGACTTAATTTCACCTCAAG 180

621 cggagaaacgcgacccgacccagcccccgcgacgacgaccccaaacccacac 890  
 |||||  
 181 TGGAGATCCCTGACCTTGAACACAGCCCTTGACAGCTCTGCCCCCTCAACCTCACC 240  
 Db

241 CTGACCTCCGCTGCTATGAGCTACTGCACTACCTCAAGGCATATGCAAGTTGTGCC 300

Db 301 CTGCACCAATTACACTGAACTAGAGGGAGTTGGCAGTGGCGTATGAAAAACATT 360

QY	1070	-ggcctaatgycacacatttacttgcatttgccttgaagtgagtaacgatttttttttc	112
Db	361	GAACAGTTTCTCGATGGCCTAGCTCCCTTATAAACCAAGAGCCTTCAGACCCCTTACAAA	420

Db 421 GGCTTAATGGGACATTTTACTTGCATTGCTTGGAGTAGTAAAGCGTTTTTTTTTC 480

Qy 1129 tctaagaanaatcgacgcttctttttaaatgcgtacttatgga 1175  
|||||  
Db 481 CCTAAGAAATCGCAGGCTTC-TTTTTTAAATGCTGCCTTTATGGA 526

RESULT	4			
LOCUS	BF692596	559 bp	mrna	EST
DEFINITION	602248949P1 NIH_MGC_62		homo sapiens cdna clone	IMAGE:4334100 5'
ACCESSION	BF692596		mrna sequence.	
VERSION	BF692596.1		GI:11978004	
KEYWORDS	EST.			
SOURCE	human.			

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 559)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Struhsaker, Ph.D.

Tissue Procurement: ATCC/DCTD/DTP  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution through the I.M.A.G.E. Consortium

```

http://image.lnl.gov
plate: L10M1203 row: a column: 13
High quality sequence stop: 555.
location/Qualifiers
1..559
source
    organism="Homo sapiens"
    /organism="Homo sapiens"

```

```
/clone_lib="NRH_MGC_62"  
/tissue_type="melanotic melanoma, high MDR"  
/lab_host="DH10B (T1 phage-resistant)"
```

SfilI (ggccgcctgagcc): Site\_2: SfilI (ggccatgatgccc):  
Double-stranded cDNA was prepared from cell line RNA.  
and 3' adaptors were used in cloning as follows: 5'

sequence: 5'-ATTCTAGAGCGCCGAGCGCCGAGATG-(dr)(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies

full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

Query Match	41.1%	Score 482.8;	DB 11;	Length 559;
Best Local Similarity	96.8%	Pred. No. 1.8e-112;		

394 tatgattccgagatggtcttatcatgtgtgatatgaaatgacatcagaagctttaatgtg 45

Qy 454 cgcattatcgagaccatatltgtagtltgtacgaagcctgtlltglttcggaccacc 51  
|||||  
61 cccgattatcgagaccatta-tgcacgctttacttcdaagcccttc-ctttcttccccgagacc 11

QY 514 atcagtaacaatacaagtatgagcagagatgctgactgatagaagatgagagagca 57  
|||||  
Db 119 ATCAGTAAACAATATCAAGTATGACGAGANGATGTGACTGATGAAGACAGATGGACAGAGCA 17

QY 574 gcaagsgaagcaaatgcgatgatttatacatgaggttccataataatttaabacattg 63

```

Db 179 GCAAGGACCAATGATGATTTATCATGAGCTTCCATAAATTAATATGATG 238
Qy 634 gttaggggaaaa-aggagctcaatgagtgagggcagaacaagagatcgcaattgctcg 692
Db 239 GTAGGGGAAAACAGAGACTCAATGATGAGGCGACAAACAGAGATCGCAATTGCTCG 298
Qy 693 tgccttagtcgaaacccaagattcgtatttagatgaggtacgtctgcctgagatc 752
Db 299 TGCCTTAGTTCGAAACCCCAAGATTCTGATTTAGATGAGCGCTACGTGCGCTGATTC 358
Qy 753 agaaagaatcagctgtttaaagctgacgtggagaaagataccccaggtatttatttg 812
Db 359 AGAAAGGAGTACAGCTTCTCAAGCTGAGAGAGATACCCCAAGTATTCATTTTG 418
Qy 813 acctaatcaccctcaagtgagatcgctgacattgaaccaagcccttcgacagctct 872
Db 419 ACCGAATTTCACCTCAGTGGAGATGCGTACCTTGACACAGGCCCTTGACAGCTCT 478
Qy 873 ggcctcctcaaacctcaccctgacctctgctgctactgagctactgacatacctcaagg 932
Db 479 GGCCTCCCAAC--TCACCTGACCTCCTCGTACTGACTATGAGTACTGACATCAAG 536
Qy 933 ccaatgcaagttgtgcccctgac 956
Db 537 -CATATGACGTTGTGGCTTGAC 559

RESULT 5
AA243820 405 bp mRNA EST 06-AUG-1997
LOCUS 2167906.r1 Soares_NhMPu_S1 Homo sapiens cDNA clone IMAGE:668506 5'
DEFINITION similar to SW:MDRI_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;
mRNA sequence.
AA243820
VERSION AA243820.1 GI:1874631
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
THIS clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1572 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham.
FEATURES
source
1. 405
/organism="Homo sapiens"
/db_xref="GDB:5562473"
/db_xref="taxon:9606"
/clone="IMAGE:668506"
/clone_1lb="Soares_NhMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2Nbm, pregnant uterus
NbHPU, and fetal heart NbH19m) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA

```

```

BASE COUNT 115 a 81 c 85 g 124 t
ORIGIN
Query Match 23.8%; Score 279.6; DB 10; Length 405;
Best Local Similarity 98.6%; Pred. No. 1e-60;
Matches 282; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
340488-345479, and 484488-489479."

Qy 160 tccctcattatgtaagaacccagatagataactttccacagctgatatcaact 219
Db 120 TTCAGGTTATTGATTAAGAAACCCAGTATAGATTAACCTTTCACAGCTGATTAACCT 179
Qy 220 gaatccatagaagaagactggtgaaatttaaaatttcttcaatttccatcaagaaca 279
Db 180 GAATCCATAGAGAGAACTGTGAAATTTAAATGTTCTTCAATATTCATCAAGACCA 239
Qy 280 tctatcaagattcgaagagctgatactcagaattagctgagagacagctgcttg 339
Db 240 TCTATCAAGATTCTGAAGAGCTGATCTGATCTGAGAGACAGTGCCTTG 299
Qy 340 gtccgtcctcaatgagcagtggaagagtaagctagctcagctctgacagaggtatagat 399
Db 300 GTCGCTCATATGACGACAGGAGAGAGTACGGTACTCCAGCTTTCAGAGAGGTTATGAT 359
Qy 400 ccgagatgagcttatcatatgctgtagatgagatgacatcaagatt 445
Db 360 CCGGATGATGGCTTATCATGTGTGATGAGATGACATCAGAGCTT 405

RESULT 6
BF584668 894 bp mRNA EST 12-DEC-2000
LOCUS 602098406.F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
DEFINITION mRNA sequence.
ACCESSION BF584668
VERSION BF584668.1 GI:11658386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LLM9798 row: 1 column: 02
High quality sequence stop: 651.
FEATURES
source
1. 894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4218385"
/clone_1lb="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

```



OY	683	caatgcctcgtcccttagtgcgaaccccccaagattcatttgataagtcactgcty	742
Dd	589	CAATAGCGAAGGCGCAATGCTAAAAAAATCCAGCATTCTTCTCCTGCATGACGGACAAGTGT	648
OY	743	cacctgatccagaagaagcaagtccagtctgtccaagtcacgtccactggta	785
Dd	649	CATTGGACTCTGAATCAGAAAAGCTGGTGCAAGAAGCACTTA 691	
RESULT	8		
AL520322			
LOCUS			
DEFINITION	ALS20322 LTR1.NFL004.NBC2 Homo sapiens cDNA clone CSDB006YC15 5 prime mRNA sequence.	EST	13-FEB-2001
ACCESSION	ALS20322		
VERSION	ALS20322.1	GI:12783815	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 943)		
TITLE	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage Bp 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr. Location/Qualifiers		
FEATURES	source		
	1..943		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CSDB006YC15"		
	/clone_1lb="LTR1.NFL004.NBC2"		
	/sex="male"		
	/tissue_type="neuroblastoma cells"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	253 a 225 c 253 g 211 t	1 others	
ORIGIN			
Query Match	18.8%; Score 220.6; DB 10; Length 943;		
Best Local Similarity	58.6%; Pred.No.1.3e-45;		
Matches 404:	Conservative 0; Mismatches 279; Indels 6; Gaps 1;		
OY	108	ctgtagtttctttagtgaatcatcatagcatttcatgttgagcagcagtcctcat	167
Dd	52	CAGTTTCATTGGCTCTCGACTATGCCAAGCCAAAATATCAGCAGCCCATCATGATGAT	111
OY	168	tattgataagaaccagatatagataactttccacagctgtgatatataaacctgaatcat	227
Dd	112	CATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGGAAGCCTTAATGCCAGACACATT	171
OY	228	aagaagaactcgtggaatttaaaatglttcttcaattcatcatcaagaaccatcataca	287
Dd	172	GGAAGGAATATGCACATTGGTGGAAGTGTATTCACATATCCACCACGACGGACATCCC	231
OY	288	gattctgaaaggctcgaatcagaatcagaattaagtcgtggaagacagtcgccgttgtcgt	347
Dd	232	ACTGCTTTCAGGAGACTGAGCCTGGAGGTGAAGAAGGCCAGACGCTGCTGTGGGCGAG	291
OY	348	caatgycagctggaagaagtaagtgatgccagcttcgtcgaaggttatatgatccgatat	407

Db	292	CAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCTGGAGCGTTTCACGACCCCTTGGC	351
Oy	408	tggtcttatcattggtgataagatgacatcaagagctttaagtgtcgcaattatcgaga	467
Db	352	AGGGAAGTGGCTGCTTGGATGGCAAGAAATAAAGCGACTGATGTTCAGTGGCTCGAGC	411
Oy	468	ccatatgtgagtggtgtatgaatcaagagccgttttgttcggagccaccatcagtaacaatat	527
Db	412	ACACCTGGGCAATCGTGTCCACGAGACCCCATCTCTTTTACTGTGCAGCATTTGCTGGAACAT	471
Oy	528	caagatctga-----cgagatgatactgcatgataagagatgagagagagaagaga	581
Db	472	TGCCATTGAGAGACACACACCCGGGTGTGTTCACAGSANAAGATCGTAGGCGCAGCAAGGA	531
Oy	562	agcaaatgcgtatgatttatacattgagatttccttaataatatacatatggtagaagga	641
Db	532	GGCCCAACTATCACTGCTTCATCGAGTCACTGCTCTTAATAATATGACATAAAGTAGAGGA	591
Oy	642	aaaaggaactcaaatgaatgagtagagggcagaagaagagatgcgaattgctgtgcttaagt	701
Db	552	CAAAAGAACTACTGCTCTCTGTGGGCAAAACACACGCTTGCATAGCTGTGCCCTTGT	651
Oy	702	tgcgaaccaccaagattcgcattttagatcagagctacgctgccttgatataagaagcaa	761
Db	652	TAGACAGGCTCATATTGCTTGGATGAGAAGCCACGTACAGCTGTGATACAGAAAGTGA	711
Oy	762	gtcagctgttcaagctgcactggagaag	790
Db	712	AAAGTTGTCCMAAGACCTGGACMAAG	740
RESULT	9		
LOCUS	AZ540627	886 bp	GSS
DEFINITION	ENTEO18TF Entamoeba histolytica Sheared DNA	Entamoeba histolytica	14-NOV-2000
ACCESSION	AZ540627		
VERSION	AZ540627.1	GI:11147603	
KEYWORDS	GSS.		
SOURCE	Entamoeba histolytica.		
ORGANISM	Eukaryota; Entamoebidae; Entamoeba.		
REFERENCE	1 (bases 1 to 886)		
AUTHORS	Loftus,B., Van Aken,S. and Fraser,C.		
TITLE	Determination of clone end sequences from Entamoeba histolytica		
JOURNAL	HMI:IMSS sheared DNA library		
COMMENT	unpublished (2000)		
	Contact: Brendan J Loftus		
	Department of Eukaryotic Genomics		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel: 301 838 0200		
	Fax: 301 838 0208		
	Email: entae@ligr.org		
	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared		
	DNA library		
	Seq primer: M13-Forward		
	Class: shotgun		
	High quality sequence start: 23		
	High quality sequence stop: 856.		
FEATURES			
Source	Location/Qualifiers		
	1..886		
	/organism="Entamoeba histolytica"		
	/strain="HMI:IMSS"		
	/db_xref="taxon:5759"		
	/clone_lib="Entamoeba histolytica Sheared DNA"		
	/note="Vector: pHOST1, Site 1; Bst I; Constructed at The		
	Institute for Genomic Research (TIGR), Rockville, MD.		
	Genomic DNA isolated from broth cultures of E. histolytica		
	using a method described by Clark and Diamond (Clark,		
	C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a		
	method for isolate identification. Exp. Parasitol.		



Db	353	ATGGCAATGACATGATTTATTTTCATCTCTCCAGAGATATGATACATGTTAGTA	294
Qy	640	gaaaaagagctcaaatgagtggagggacgaacagagatcgcaattgctcgcccta	699
Db	293	GAAGAAAGAGCTGCATTTTCAGAGGAGCAAAAAACAAAGATTTGTCATTCATTC	234
Qy	700	gttcgaaccccaagatctcgatttagatgagtgctgccttcggtatcgaagc	759
Db	233	ATTAGAAACCTTCACATTTTATTACTTATGAGATCATCATGACCTTGATACACAAAGT	174
Qy	760	aagtcagctgtcaagctcgaactggagaagata	793
Db	173	GAAAGATTGTACACACAGCAGCTGAAAAAGCTA	140
RESULT	11		
LOCUS	A2687805/c		
DEFINITION	A2687805	880 bp	DNA
ACCESSION	ENTLUS27F	Entamoeba histolytica	Sheared DNA
VERSION	A2687805		GSS
KEYWORDS	A2687805.1	GI:11824951	14-DEC-2000
SOURCE	GSS.		
ORGANISM	Entamoeba histolytica.		
REFERENCE	Entamoeba histolytica.		
AUTHORS	Eukaryota; Entamoebidae; Entamoeba.		
TITLE	1 (bases 1 to 880)		
COMMENT	Loftus, B., Van Aken, S. and Fraser, C.		
JOURNAL	Determination of clone end sequences from Entamoeba histolytica		
COMMENT	HMI:IMSS sheared DNA library		
COMMENT	Unpublished (2000)		
COMMENT	Contact: Brendan J Loftus		
COMMENT	Department of Eukaryotic Genomics		
COMMENT	The Institute for Genomic Research		
COMMENT	9712 Medical Center Dr., Rockville, MD 20850, USA		
COMMENT	Tel: 301 838 0200		
COMMENT	Fax: 301 838 0208		
COMMENT	Email: entae@igr.org		
COMMENT	Clones are derived from the Entamoeba histolytica HMI:IMSS sheared		
COMMENT	DNA library		
COMMENT	Seq primer: M13-Forward		
COMMENT	Class: shotgun		
COMMENT	High quality sequence start: 30		
COMMENT	High quality sequence stop: 858.		
FEATURES	Location/Qualifiers		
SOURCE	1..880		
	/organism="Entamoeba histolytica"		
	/strain="HMI:IMSS"		
	/db_xref="taxon:5759"		
	/clone_lib="Entamoeba histolytica Sheared DNA"		
	/note="Vector: pHOSt1; Site: 1; Bst I; Constructed at The		
	Institute for Genomic Research (TIGR), Rockville, MD.		
	Genomic DNA isolated from broth cultures of E. histolytica		
	using a method described by Clark and Diamond (Clark,		
	C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a		
	method for isolate identification. Exp. Parasitol.		
	77:450.). The DNA was mechanically sheared to give a		
	tight size distribution (~2 kb). The v + i method used for		
	the library construction is described in detail in Smith,		
	H.O. and Venter, J.C. (Making small insert libraries for		
	whole genome shotgun sequencing projects. In Genome		
	Sequencing: A Practical Approach, eds. M. Vaudin and B.		
	Barell, Oxford University Press, 1999)."		
BASE COUNT	247 a	167 c	122 g
ORIGIN	344 t		
Query Match	17.7%;	Score 207.6;	DB 13; Length 880;
Best Local Similarity	59.3%;	Pred. NO. 2.6e-42;	
Matches 376;	Conservative 0;	Mismatches 24;	Indels 9; Gaps 1;
169 attgataagaaaccagatataactttccacagctgataataaccctgaatccata	228		

FEATURES	LOCATION/QUALIFIERS	SEQUENCE	START	END	ORIENTATION	FEATURE	QUALITY
Db	198	GAAGAAGATTGTACACACAGCACTTGA	165				
Qy	760	aagtcagctgtccaagctgcactgtagaagata	793				
Qy	700	gttcgaacccccaaagcttcgtatttagatgagctcagctcgtcgtgataagaagc	759				
Db	258	ATTGAAAACTTCACTTATTTACTTGATGAAGCTACATCAGCACTTGATGATACCAAACT	199				
Qy	409	ggcttatcatcgttgatgataagaaatgacalcagagcttaaatgtcgcgcatlctgagac	468				
Qy	469	catattgtagtggttaagcttaagaagcctgttctgttcggagccacatcagtaacaaatc	528				
Db	498	CAAAATAGGATTTAGTTGGACAGAACCTATTTGTTGTCATGTACATTTATGAGAGAAATATTT	439				
Qy	529	aagttatgga-----cgagatgactgcactgaagagagatgagagagagcaagc	579				
Db	438	ATGCTTTGGAGCTAGAGATGAGACAAACCACTGAAGAAAGATGATTTGAATGTGCTAA	379				
Qy	580	gaagcaaatgctgatalgatattlcatcagtagcttccctaataatltacatcgttagag	639				
Db	378	ATGGCAAAATGCACATGATATTTATTTCTCATCTTCAGAGAGATATGATCAATGATGAGA	319				
Qy	640	gaaaaaagagctccaatgtagtgagggcagaaacagagatgcatctgtcgtccta	699				
Db	318	GAAGAAAGGAGCTGCATTATCAGAGAGACAAAACAAAGAAATTCATTTGACGCTGATTG	259				
Qy	700	gttcgaacccccaaagcttcgtatttagatgagctcagctcgtcgtgataagaagc	759				
Db	258	ATTGAAAACTTCACTTATTTACTTGATGAAGCTACATCAGCACTTGATGATACCAAACT	199				
Qy	760	aagtcagctgtccaagctgcactgtagaagata	793				
Db	198	GAAGAAGATTGTACACACAGCACTTGA	165				
RESULT	12						
LOCUS	AZ670821/C	932 bp	DNA	GSS	14-DEC-2000		
DEFINITION	ENTJN697F Entamoeba histolytica Sheared DNA	Entamoeba histolytica					
ACCESSION	AZ670821						
VERSION	AZ670821.1	GI:11807967					
KEYWORDS	GSS.						
SOURCE	Entamoeba histolytica.						
ORGANISM	Entamoeba histolytica.						
REFERENCE	Loftus,B., Van Aken,S. and Fraser,C.						
AUTHORS	Determination of clone end sequences from Entamoeba histolytica						
TITLE	HML-IMSS sheared DNA library						
COMMENT	Unpublished (2000)						
JOURNAL	Contact: Brendan J Loftus						
COMMENT	Department of Eukaryotic Genomics						
COMMENT	The Institute for Genomic Research						
COMMENT	9712 Medical Center Dr., Rockville, MD 20850, USA						
COMMENT	Tel: 301 838 0200						
COMMENT	Fax: 301 838 0208						
COMMENT	Email: entae@ligr.org						
COMMENT	Clones are derived from the Entamoeba histolytica HML-IMSS sheared						
COMMENT	DNA library						
COMMENT	Seq primer: M13-Forward						
COMMENT	Class: Shotgun						
COMMENT	High quality sequence start: 46						
COMMENT	High quality sequence stop: 867.						
COMMENT	Location/Qualifiers						







Db	415	ACGATTTTGC	AAGAAAGAGTCTGAAAGTCCCTCAGCGAAAGACTATTGCTTTGGTCGGA	356
Oy	346	ctcaattggcagtggaagagatgagtgatgctcagcttctgcagaggtatataatgacgcgat	405	
Db	355	GCCAGTGGCTGTGGGAAAAGCACACGACATTCAGTTTCGGACGCCCTTTTACGACCATGAT	296	
Oy	406	gatggtttatcacggtgagatgatatgacatcacagcagctttaaaltgtagcattatcga	465	
Db	295	TCTGTGAGAGGTGACTGTGTGATGTGTCATGATATTCGCGTCCCTTAATGTGTGCGCTGAGG	236	
Oy	466	gaccatattgagatggttagtcaagaagcctgttltgttcggaccacacatagtaacaat	525	
Db	235	GAGAACATGCGGGGATTTGTGTAGTCAGGAACCACTGCTTTTGTGAACCCACCATTTGCAGAGAAC	176	
Oy	526	atcaagaatgtagcagagatgatatgctatcaataagaatgtagagagagcagaaggaaaca	585	
Db	175	ATTGCTTATGTGCGCTGGAAGACGCCACCGATTCAGATATTGTAACAGGCCATTAAAGAGGCC	116	
Oy	586	aatgcgtatgattttatcatgattgagatttctctaataattaatacattgtagtgggaaaaa	645	
Db	115	AATGGTATCGATTTTCATTTCCAACTTCCAGACGACATACACATATGTTGTGTAAGAGA	56	
Oy	646	ggagctcaaatgagtgtagggcgagaacagagatgcgaattgctgtgccttag	700	
Db	55	GGTGCTGAGATTAACGGAGGACAAAAACAGAGGATGCGCATCGCTGGGGCTATGG	1	
RESULT	14			
LOCUS	AM342644/C			
DEFINITION	AM342644 575 bp mRNA EST 31-JAN-2000			
	181106.x1 Sugano Kawakami zebrafish DRB Danio rerio cDNA clone			
	2644067.3 similar to TR:Q91586 Q91586 MULTIDRUG RESISTANCE			
	PROTEIN. ; mRNA sequence.			
ACCESSION	AM342644			
VERSION	AM342644.1			
KEYWORDS	GI:6839010			
SOURCE	EST.			
ORGANISM	zebrafish.			
	Danio rerio.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;			
	Cypriniformes; Cyprinidae; Rasbora; Danio.			
AUTHORS	1 (bases 1 to 575)			
	Sugano S., Kawakami K., Johnson S., Li F., Marra M., Eddy S.,			
	Hillier L., Clifton S., Allen M., Gibson M., Jost S., Kucaba T.,			
	Matlin J., Page D., Steptoe M., Underwood K., Theisling B., Rutter			
	E., Bowers Y., Wylie T., Waterston R. and Wilson R.			
	Washu zebrafish EST Project 1999			
	Unpublished (1999)			
TITLE	Contact: S.L. Johnson			
JOURNAL	Washington University School of Medicine			
COMMENT	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA			
	Tel: 314 286 1800			
	Fax: 314 286 1810			
	Email: est@watson.wustl.edu			
	Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA			
	Sequencing by: Washington University Genome Sequencing Center			
	Seq primer: T7 RT from Amersham			
	High quality sequence stop: 465.			
FEATURES	Location/Qualifiers			
source	1..575			
	/organism="Danio rerio"			
	/strain="AB"			
	/db_xref="taxon:7955"			
	/clone="2644067"			
	/clone_lib="Sugano Kawakami zebrafish DRB"			
	/sex="mixed (one male and one female, including			
	unfertilized eggs)"			
	/dev_stage="adult"			
	/lab_host="DH10B (phage resistant)"			
	/note="Vector: pME185-FUS; Site_1: DraIII (CACTGTG);			
	/note_2: DraIII (CACCATGG); 1st strand cDNA was primed			
	with an oligo(dT) primer [ATGCGCCCTTTTTTTTTTTTTTTT];			

BASE COUNT	134 a	148 c	129 g	164 t		
ORIGIN						
Query Match	17.2%; Score 201.6; DB 10; Length 575;					
Best Local Similarity	60.7%; Pred. No. 8.1e-41;					
Matches	330; Conservative	0; Mismatches	214; Indels	0; Gaps		
<p>double-stranded cDNA was ligated to a DralII adaptor (TTTGGGCTACTGG), digested and cloned into distinct DralII sites of the pME18S-FL3 vector (5' site CACGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments &lt;1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science) and kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTCCTGCTTAAACCTGCG and 3' end primer CGACCTGCGAGCTGACGACA. "</p>						
OY	131	ccatagcagattatcattgtagcagcagccctcatattgtataagaaccagatag	190			
Db	544	CCAAAGTCGGCGCGCTGCTCTCAATAGGTTGTCATATACATGACCTCTCCATTG	485			
OY	191	ataactttccacagctgatatataaccgtgaatccatagagaactcgtggaattaa	250			
Db	484	ACAGCAGTTTCCAAAGAGACACAAACCGCCGTGTGAGAGAGATTCGAATTCAAA	425			
OY	251	atgttcttccaattatccatcaagaaccatctataagaattctgaaaggtctgaatc	310			
Db	424	ACATTATATTTCAATTACCCCTCAGAAAAGATGTGACGATTTTGCAGGAATGAGTCA	365			
OY	311	gaattaagtctgagagagaagatcgctcgttggtgtgtctcaatgagcagtggaagatg	370			
Db	364	AAGTCCCTCACGAAAGACTATTGCTTGTGTGAGACAGTGCTGTGGAAAAGCACCA	305			
OY	371	tagtcacgctctgagaaggtatataatgcagatgatgtgtcttcatcagtggtgatga	430			
Db	304	CGATTGAGTGTGTGACGCGCTTTTACGACCTCGATTCGTGAGAGGAGACTGTGAGTGC	245			
OY	431	atgacatacagagctttaaattgtcggcattatcgagaccatattggagtgttatgtaag	490			
Db	244	ATGATATCCGCTCCCTTAATGTGCGGTGCTGAGGAGAAATGGGATGTGAGTCAGG	185			
OY	491	agccgttttctgtcggagcaccatcagtaacaatatacaagtataagcagagatgata	550			
Db	184	AACCACTTTCTTTTGGAAACCACTATTCGACAGAACACTTGGTATATGSCCTGAAGCGCA	125			
OY	551	ctgatgaagatgtagagagcagcaagggaagcaaatgctgatattatcatatgagat	610			
Db	124	CCGATGCAATATTCGAAAGCGCATTTAAAGAGGCCAATCTCAACGATTTTCATTCCAAAC	65			
OY	611	ttccctaataattatacatattgtagaggaagaaagagactcaaatgagtgtaggcaga	670			
Db	64	TTTCCAGATTAACGAACACTATGTGTGTGAGAGAGAGTCTCAGCTTAACGCGAGGACAGA	5			
OY	671	aaca	674			
Db	4	AACA	1			
RESULT	15					
H82550	402 bp	mrna	EST	09-NOV-1995		
LOCUS	YV77A07.r1	Scores melanocyte 2NBHM Homo sapiens cDNA clone				
DEFINITION	IMAGE:248724.5	similar to SP:MDRI-CHTR P21448 MULTIDRUG				
ACCESSION	H82550	RESISTANCE PROTEIN 1 ;, mRNA sequence.				
VERSION	H82550.1	GI:1060639				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 402)					



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 23:02:05 ; Search time 886.72 seconds

(without alignments)  
1136.049 Million cell updates/sec

Title: US-09-873-409-11

Perfect score: 1175

Sequence: 1 cccaatccctaatatc.....taaaatgctgactttatgga 1175

Scoring table: IDENTITY-NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N.Geneseq\_1101.\*

- 1: /SIDS2/gcgdata/geneseq/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324.2	27.6	4279	22	AAD03488 Dog P-glycoprotein
2	324.2	27.6	4279	22	AAD03504 Dog P-glycoprotein
3	324.2	27.6	4279	22	AAD03505 Dog P-glycoprotein
4	324.2	27.6	4279	22	AAD03506 Dog P-glycoprotein
5	322.6	27.5	4317	22	AAD03489 Dog P-glycoprotein
6	319.4	27.2	3988	21	AAZ88973 Human MDR-1 DNA.
7	319.4	27.2	4669	14	AAQ52726 Sequence of human
8	317.8	27.0	3860	21	AAZ49332 Human wild-type mu
9	317.8	27.0	3860	21	AAZ49333 Human G185V mutant
10	317.8	27.0	4264	19	AAV6533 Mutated human P-g1
11	317.8	27.0	4264	19	AAV6534 Mutated human P-g1

12	317.8	27.0	4349	22	AAH57442
13	317.8	27.0	4378	11	AAQ04522
14	317.8	27.0	4646	15	AAQ72872
15	317.8	27.0	4646	15	AAZ94738
16	317.8	27.0	4669	8	AAV70752
17	317.8	27.0	4669	19	AAV32645
18	317.8	27.0	6505	17	AAV13394
19	317.8	27.0	8630	21	AAZ13041
20	317.8	27.0	8630	21	AAZ24042
21	314.4	26.8	4788	21	AAZ49335
22	311.2	26.5	4369	21	AAZ52047
23	311.2	26.5	4425	21	AAZ52048
24	309.8	26.4	4186	22	AAZ86127
25	309.8	26.4	4195	22	AAZ86128
26	309.6	26.3	4233	21	AAZ90198
27	309.6	26.3	4233	22	AAZ90198
28	304.8	25.9	4189	21	AAZ49334
29	304.8	25.9	4313	14	AAQ38950
30	287.4	24.5	3924	21	AAZ94742
31	287.4	24.5	3924	21	AAZ94742
32	245.8	20.9	4776	21	AAZ94744
33	220.6	18.8	2726	15	AAQ70907
34	220.6	18.8	2726	15	AAQ70916
35	220.6	18.8	2726	18	AAZ43322
36	205.4	17.5	465	22	AAI38132
37	204	17.4	216	22	AAI51285
38	199	16.9	4102	21	AAZ45942
39	184.8	15.7	14000	21	AAQ09020
40	183.2	15.6	4051	21	AAA09019
41	182.6	15.5	1957	21	AAZ35415
42	179.6	15.3	3512	20	AAV69395
43	177.8	15.1	6300	21	AAV60700
44	176.4	15.0	1810	20	AAV69393
45	176.4	15.0	4175	20	AAV69392

#### ALIGNMENTS

RESULT 1	
AAD03488	standard; cDNA; 4279 BP.
ID	AAD03488
AC	AAD03488;
XX	
XX	13-JUN-2001 (first entry)
DT	
XX	
DE	Dog P-glycoprotein (PGP) cDNA #1.
XX	
XX	Dog: P-glycoprotein; PGP; multidrug transporter; MDR1;
KW	drug bioavailability; transgenic animal; genetic model; ss.
XX	
OS	Canis familiaris.
XX	
XX	
FT	Key
FT	Location/Qualifiers
FT	17.3862
FT	/tag="a
FT	/product="Dog P-glycoprotein (PGP) #1"
XX	
XX	WO200123540-A2.
XX	
XX	05-APR-2001.
XX	
XX	28-SEP-2000; 2000WO-US26767.
PF	
XX	
XX	28-SEP-1999; 99US-0156510.
PR	
XX	
XX	(GENT-) GENTEST CORP.
PA	
XX	
XX	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
PI	WPI, 2001-235373/24.
XX	
DR	P-PSDB; AAE00303.

Human intestine ce  
Multidrug Resistan  
Human multidrug re  
Human ATP binding  
Human P glycoprote  
Sequence of human  
Hybrid vector pSF  
Retroviral M4 mdr-  
Retroviral vector  
Murine multidrug r  
Rat multidrug resi  
Rat multidrug resi  
Cynomolgous monke  
Rat mdr1b2 multir  
Rat mdr1b2 multir  
Murine multidrug r  
Mouse multidrug re  
Human ATP binding  
Human MDR-3 DNA.  
Human ATP binding  
Multidrug-resistan  
Multidrug-resistan  
Multidrug resistanc  
Probe #6818 used t  
Probe #19971 used  
Arabidopsis thalia  
A. thaliana ATPAC  
A. thaliana ATPAC  
Arabidopsis thalia  
H. contortus PGP-O  
ACPGP-1 ATP bindin  
H. contortus PGP-A  
H. contortus PGP-A

XX New dog P-glycoproteins (Pgp) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT Pgp inhibitors -  
XX  
PS Claim 3; Page 58-63; 11pp; English.

Qy	766	gctgttcaagctgcaactgagaaag	790
Db	1724	gtgtctcagltgcccctgataaag	1748

The invention relates to dog P-glycoprotein (PGP) also referred as multidrug transporter (MDR1) and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog P-glycoprotein. PGP and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterised by PGP activity, by reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.

The present sequence is dog P-glycoprotein (PGP) cDNA. This sequence is also referred as genotype C cDNA. The PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.

RESULT	2
AAD03504	
ID	AAD03504 standard; cdwA; 4279 BP.
XX	
AC	AAD03504;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Dog P-glycoprotein (PGP) allelic variant (Genotype A) cdwA.
XX	
KW	Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW	drug bioavailability; transgenic animal; genetic model; ss.
XX	
OS	Canis familiaris.
XX	
EH	Location/Qualifiers
FT	17..3862
FT	/*tag= a
FT	/product= "Dog P-glycoprotein (PGP) allelic variant
FT	(Genotype A) protein"
FT	replace (607, C)
FT	allele

Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Query Match	27.6%	Score 324.2	DB 22	Length 4279
Best local Similarity	69.9%	Pred.No. 1.4e-82		
Matches 437; Conservative	0	Mismatches 188	Indels 0	Gaps 0

Oy	166	atatttgtaagaacccadatatagataaacatttcacagccggtatgataaaccttaacc	225
Db	1124	ataattgcataaacaacgaacatttgacaactatctgaaagatgacataaaaccgataat	1183
Oy	226	atagaagaactctgtggaatttaaaatgltttcttcaattatocatacgaaccatctac	285
Db	1184	attaagagaaaattctggaattcaaaaatglttcattcattcgaattacccttctcgaaaagaatt	1243
Oy	286	aagattctgaaaagttctgaaactcagaattaaatctcggagagacaaatgccttgctgct	345
Db	1244	aagacttaaaagggtctcaacccgtaaggttcgaagtctggagacagctggcgtgtgtgg	1303
Oy	346	ctcaatgtgcagctggagaaagatcagctgtaagtcacgctcttcgacaggttatatccgcat	405
Db	1304	aacagctgctggagagagacgaccggtgagctgtagagaggtcttatgacccacaa	1363
Oy	406	gatggtcttcatcagtggtgagatgagaatgacatcagaagctttaaattgtgagcatattcga	465
Db	1364	gattgacatcagtgctgtatgatagagacaagacattagacaataatgtlaagagacttcogg	1423
Oy	466	gaccatcttggaagtgtatgacgaagaagccgtttgtctggagacacacatcagtaacat	525
Db	1424	gsaattactggttggtgtagcagagacccgtgtgtttgcccacacagctagctgaaac	1483
Oy	526	atcaaagtatggacgaagatgactgcatgaaagagatgagagagcagcaaggaagca	585
Db	1484	atctgcattgcccggaaaatgltacacatgtagatgattgagaagctgttlaagaaagcc	1543
Oy	586	aattgctatgattatcatcacaaggattccctcaataaattatcatcattgttgtagggaaaa	645
Db	1544	aattgcctatgattatcatcacaaggattccctcaataaattatcatcattgttgtagggaaaa	1603
Oy	646	ggagctccaatgagtgtgaggggcagaaacagagatcgcgaattctgcgtgcttaagtcca	705
Db	1604	ggggcccaactgaatggtgtgacacaaaacagaaatcgcgaattctgcgtgcccgtgtgc	1663

Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other:

QY	Db	Query Match	Score	DB 22	Length
706	aaccacaaatctctgattcttaga	27.68	324.2	4279	705
1664	aaccacaaatctctgattcttaga	60.98	1.4e-82	0	705
	aaccacaaatctctgattcttaga	437	188	0	705

```

OY 166 attattgataagaaccagatagataactttccacagctgagatataaactgatacc 225
DB 1124 ataattgacaataaaccagcatctgacagctatccgaagaatgagacataaaccagataat 1183
OY 226 atagaaggaactgtggaatttaaaaaatgtctcttcaatlatcacaagaccatlatc 285
DB 1184 ataaagggaatttggaattcaaaaatgttcaacttcaacttccctctctgaaagaagtt 1243
OY 286 aagatctgaaagctcgaatcagaatlaagctctgagagacagctgcctgtgcgt 345
DB 1244 aagatcttaagagctcacaacctgaagctcagagtgaggcagacagctgagctgtg 1303
OY 346 cccaatgggaatgggaagagctgagctcagctccttcagagctatagctcggat 405
DB 1304 aacagatgctgcgggaagaacagcagcagctgagatgagagagctctcatgacccaca 1363
OY 406 gatgcttcatcatgctgagatgagaatgacatcagacgttlaaatgtcgcatatcga 465
DB 1364 gatgcatgctgctgatttgatgacagagcatatgagacataatgtaagcctctcgg 1423
OY 466 gacacattggaatgtgtgtcgaagagcctgtttgttcggagacacacacagtaacaat 525
DB 1424 gaattactggtgtgtgtgagtcagagagcctgtgtgttgcaccacagatagctgaaac 1483
OY 526 atcaagtatgagcagatgagatgactgataagagatgagagagcagcagaaggaagca 585
DB 1484 attcgctatgagcgcggaataatgcaccatgagatgagagatgagaacgttgaagagac 1543
OY 586 aatgcgtatgatttatacctgagagcttctcctaataatlaatacatgtgtagggaaaaa 645
DB 1544 aatgcctatgatttatacctgagagcttctcctaataatlaatacatgtgtagggagaga 1603
OY 646 ggaagctcaaatgagtgagagggcagaacagagatgcgaatgtcgtcgttaagtcga 705
DB 1604 gggggcccaagtgtgtgtgagacagaaacagagatgcgaatgtcgtcgtcgtgtcgc 1663
OY 706 aaccccaagaactgatttataagatgagctgagctgccttgatccagaagaacagtcga 755
DB 1664 aaccccaagaacttctctgctgagtgagagcagctgccttgatccagaagaatgtaga 1723
OY 766 gctgttcaagctgacatggaagaag 790
DB 1724 gtggttcaagtgagccttgatgaag 1748

RESULT 3
AAD03505
ID AAD03505 standard; cDNA: 4279 BP.
XX
AC AAD03505;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KM Dog: P-glycoprotein allelic variant; PGP: multidrug transporter;
XX MDR1: drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT /tag= b
FT /tag= c
FT allele replace (607, C)
FT
XX
XX WO200123540-A2.
XX

```

```

PD 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steinel-crespi DT, Crespi CL, Reif TC, Pattem CT;
XX WPI; 2001-235373/24.
XX P-PSDB; AAE00309.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors.
XX
XX Claim 9; Page 93-99; 11pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Query Match 27.6%; Score 324.2; DB 22; Length 4279;
Best Local Similarity 69.9%; Pred. No. 1,4e-82;
Matches 437; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

OY 166 attattgataagaaccagatagataactttccacagctgagatataaactgatacc 225
DB 1124 ataattgacaataaaccagcatctgacagctatccgaagaatgagacataaaccagataat 1183
OY 226 atagaaggaactgtggaatttaaaaaatgtctcttcaatlatcacaagaccatlatc 285
DB 1184 ataaagggaatttggaattcaaaaatgttcaacttcaacttccctctctgaaagaagtt 1243
OY 286 aagatctgaaagctcgaatcagaatlaagctctgagagacagctgcctgtgcgt 345
DB 1244 aagatcttaagagctcacaacctgaagctcagagtgaggcagacagctgagctgtg 1303
OY 346 cccaatgggaatgggaagagctgagctccttcagagctatagctcggat 405
DB 1304 aacagatgctgcgggaagaacagcagcagctgagatgagagagctctcatgacccaca 1363
OY 406 gatgcttcatcatgctgagatgagaatgacatcagacgttlaaatgtcgcatatcga 465
DB 1364 gatgcatgctgctgatttgatgacagagcatatgagacataatgtaagcctctcgg 1423
OY 466 gacacattggaatgtgtgtcgaagagcctgtttgttcggagacacacacagtaacaat 525
DB 1424 gaattactggtgtgtgtgagtcagagagcctgtgtgttgcaccacagatagctgaaac 1483
OY 526 atcaagtatgagcagatgagatgactgataagagatgagagagcagcagaaggaagca 585
DB 1484 attcgctatgagcgcggaataatgcaccatgagatgagagatgagaacgttgaagagac 1543
OY 586 aatgcgtatgatttatacctgagagcttctcctaataatlaatacatgtgtagggaaaaa 645

```

```

Db      1544 aatgcctatggtttatcatcagaactaactaataattgacactctggttgagagaga 1603
QY      646 ggaactcaatgagtgagggcagaacaagagatcgcaattgtctgtgcttattcga 705
Db      1604 ggggcccaagctgagtggtgagcagaacaagaaatcgccattctggtggttctgc 1663
QY      706 aacccaagattcgtattttagatgagctacgtctgcccgtgattcagaagaagtca 765
Db      1664 aacccaagattcttctcgtctgtagtgaggaacgtcagctctgacactgaaagtgaagca 1723
QY      766 gctgttcaagctgcactgaggaagaag 790
Db      1724 gtggttcaggtggtccctgataaag 1748

RESULT  4
AAD03506
ID      AAD03506 standard; cDNA; 4279 BP.
AC      AAD03506;
DT      13-JUN-2001 (first entry)
XX
DE      Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KM      Dog: P-glycoprotein allelic variant; PGP: multidrug transporter;
KW      MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS      Canis familiaris.
XX
FH      Key Location/Qualifiers
FT      CDS 17..3862
FT      /*tag= a
FT      /product= "Dog P-glycoprotein (PGP) allelic variant
FT      (Genotype D) protein"
FT      replace (91, T)
FT      allele
FT      /*tag= b
FT      replace (607, C)
FT      allele
FT      /*tag= c
FT      replace (1001, T)
FT      allele
FT      /*tag= c
FT      replace (3458, A)
FT      allele
FT      /*tag= c

W0200123540-A2.
XX
PD      05-APR-2001.
XX
PF      28-SEP-2000; 2000WO-US26767.
XX
PR      28-SEP-1999; 99US-0156510.
XX
PA      (GENT-) GENTEST CORP.
XX
PI      Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX      WPI: 2001-235373/24.
DR      P-Psdb: AAE00310.
DR
XX
PT      New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX      for determining the bioavailability of drugs and for screening for dog
XX      PGP inhibitors -
XX
PS      Claim 9; Page 102-107; 111pp; English.
XX
CC      The invention relates to dog P-glycoprotein (PGP) also referred
XX      as multidrug transporter (MDR1) and nucleic acids encoding them.
XX      The invention also includes fragments and biologically functional
XX      variants of dog P-glycoprotein. PGP and their nucleic acids are
XX      useful for determining the bioavailability of drugs and for
XX      screening PGP inhibitors. They are useful for the diagnosis and
XX      treatment of conditions characterised by PGP activity, by
```

```

CC      reducing or increasing PGP activity in a cell. PGP nucleic acids
CC      are used as oligonucleotide probes. Complements of PGP nucleic
CC      acids are useful as antisense oligonucleotides, to induce a PGP
CC      'knockout' phenotype. They are used to prepare a non-human
CC      transgenic animal, which are valuable as genetic models for
CC      human diseases.
CC      The present sequence is dog P-glycoprotein (PGP) allelic variant
CC      (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC      exporting small molecules across the cell membrane. This enzyme
CC      is a member of the ABC transporter family.
XX
SQ      Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Query Match      27.6%; Score 324.2; DB 22; Length 4279;
Best Local Similarity 69.9%; Pred. No. 1.4e-82;
Matches 437; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY      166 attatgtaagaaccagatagataactttccacagctggatataaacctgaatcc 225
Db      1124 ataattgacaataaacacagatgtgacagctattcgaagtgagacataaacagataat 1183
QY      226 atagaaggaactgtggaattaaaatggtcttcaatcatcacaagaacatcattc 285
Db      1184 attaaggaaatttggaaattcaaaaatgttcaacttcacactcagttcccttcgaaagaagt 1243
QY      286 aagattcgaagaagctcgaatcagaattaaagctcgaagagacagctggtcggt 345
Db      1244 aagatcttaagaaggtctcaacactgaaggttcagagtgagcagagcgtggtcggtg 1303
QY      346 ctcaatggcagctgggaagagacgtagctagctcagctctcgaaggttatatgacccgat 405
Db      1304 aacagtgctgctgggaagagacgacgacgctgagctgtagcagagctctatgaccacca 1363
QY      406 gatggtctatcactggtgtagatgagaatagacatcagagcttaaatggtcgcatatcga 465
Db      1364 gatgacatgctgtlatgtagcagacagacatagagaccataaagtlaaggcatcttcgg 1423
QY      466 gaccatattgagtggttagtcaagagcctgttctgttcgggacacacatcagtaaacat 525
Db      1424 gaaattactggtgtggtgagtcagagagcctgtgtgttcgacacacagatcgagaaac 1483
QY      526 atcaagtatgacgagatgactgactgtagaagaagaatgagagacgaaggaagca 585
Db      1484 atcgctatggcgcggaanaatgltcacacatgtagatgagaaagcgtgtaaggaagcc 1543
QY      586 aatcgatgatttatcatatgaggtttcctaataattaatatgattgtaggggaaanaa 645
Db      1544 aatgcctatgatttatcatgaaacttaccataaatttgacactcgtgttggaagaga 1603
QY      646 ggaactcaatgagtgagggcagaacaagagatcgcaattgtctgtgcttattcga 705
Db      1604 ggggcccaagctgagtggtgagcagaacaagaaatcgccattctggtggttctgc 1663
QY      706 aacccaagattcgtattttagatgagctacgtctgcccgtgattcagaagaagtca 765
Db      1664 aacccaagattcttctcgtctgtagtgaggaacgtcagctctggaactgaaagtgaagca 1723
QY      766 gctgttcaagctgcactgaggaagaag 790
Db      1724 gtggttcaggtggtccctgataaag 1748

RESULT  5
AAD03489
ID      AAD03489 standard; cDNA; 4317 BP.
XX
AC      AAD03489;
DT      13-JUN-2001 (first entry)
XX
DE      Dog P-glycoprotein (PGP) cDNA #2.
XX
```

KW	Dog: P-glycoprotein; PGP; multidrug transporter; MDR1;
KW	drug bioavailability; transgenic animal; genetic model; ss.
XX	
OS	Canis familiaris.
XX	
FH	Key Location/Qualifiers
FT	CDS 70..3912
FT	/tag=a
XX	/product= "Dog P-glycoprotein (PGP) #2"
PN	MO200123540-A2.
XX	
PD	05-APR-2001.
PF	28-SEP-2000; 2000MO-US26767.
XX	
PR	28-SEP-1999; 99US-0156510.
XX	
PA	(GENT-) GENTEST CORP.
XX	
PI	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX	
DR	WPI; 2001-235373/24.
DR	P-PSDB; AAE00304.
XX	
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT	for determining the bioavailability of drugs and for screening for dog
PT	PGP inhibitors .
XX	
PS	Claim 1; Page 66-72; 11pp; English.
XX	
CC	The invention relates to dog P-glycoprotein (PGP) also referred
CC	as multidrug transporter (MDR1) and nucleic acids encoding them.
CC	The invention also includes fragments and biologically functional
CC	variants of dog P-glycoprotein. PGP and their nucleic acids are
CC	useful for determining the bioavailability of drugs and for
CC	screening PGP inhibitors. They are useful for the diagnosis and
CC	treatment of conditions characterised by PGP activity, by
CC	reducing or increasing PGP activity in a cell. PGP nucleic acids
CC	are used as oligonucleotide probes. Complements of PGP nucleic
CC	acids are useful as antisense oligonucleotides, to induce a PGP
CC	'knockout' phenotype. They are used to prepare a non-human
CC	transgenic animal, which are valuable as genetic models for
CC	human diseases.
CC	The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC	PGP enzyme functions as an efflux pump exporting small molecules
CC	across the cell membrane. This enzyme is a member of the ABC
CC	transporter family.
XX	
SO	Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other:
<hr/>	
Query Match	27.5%; Score 322.6; DB 22; Length 4317;
Best Local Similarity	69.8%; Pred. No. 3.9e-82;
Matches 436; Conservative	0; Mismatches 189; Indels 0; Gaps 0
OY	166 attatgataagaaccagtatagataactttccacagctygatatataacctgaatcc 225
Db	1174 ataattgcacaataaaccaagcattgtcacagtatttcgaagagtgagacataaacaggtaatt 1233
OY	226 atagaagaagactcgtggaaatttaaataatgttcttccaattcatccaagaaccattatcc 285
Db	1234 attaaggagaaatttggaaattccaanaattgttcacttcgaattacccttcgaaaagaagtt 1293
OY	286 aagattctgaagaagcgtcgaatctcgaaatctaagtctggagagagacagtcgcttgctcggt 345
Db	1294 aagaacttaaaaggctccaacctgaagcttcagaagtgaggcagacagctggcgtggttggg 1353
OY	346 ctcaaattggcagctgggaaagagtagcggtagtcacagctcttcacagaagttatgatccggat 405
Db	1354 aaccagtgctcggggagaaagcagcaccgtygcagctgatgacagaagctcatagccccaca 1413
OY	406 gatggcttatcatactgltgagatgaagaatgacatccagaagctttaaatgtcgcgcatattcga 465

Db	1414	gattgcacatggtcctgtatctgatgacaggaacattatgagccataatgtaaagcattcttcg	1473
Qy	466	gaccctatctggagctggtatgacaaagacctgtttgttcgcggaccacacagtaacaat	525
Db	1474	gaattactggtgtggtgagtcagagccctgtgttcttcgaaccaagatagctgaaac	1533
Qy	526	atcaagatctgacagagatgatgtactctgataagaagatgagagacagcaaggaagca	585
Db	1534	attcgtctatgcccgcgaaaaatgtcacacatgatatgatatgaaaagctgttaaggaagcc	1593
Qy	586	aatggtatggttttatcatcatgagttccctcaataaattatatacatgtgtaagggaaaa	645
Db	1594	aatgcctatgattttatcatcatgaaactacataataatttgaacactccggttggagaaga	1653
Qy	646	ggagctcaaatgagtgtgagggcagaagaacagagatctgcaattgtcgttccttaattcga	705
Db	1654	gggggcgcgcgcgagtggtggtgacagaagaagaatcgccattgctgcggccctgttcgc	1713
Qy	706	aaccccaagatcttgatctttagatgaggtcagctgcgtccctgattcagaagaagctca	765
Db	1714	aaccccaagatcttctctgtctgatatgaggaacagctcagctcgtgacactgaaatgaa	1773
Qy	766	gctgtcaagctgcacctgagagaag	790
Db	1774	gtcttcagtggtccctgataag	1798
RESULT 6			
AAZ88973			
ID	AAZ88973 standard; DNA; 3988 BP.		
XX	AAZ88973;		
XX	30-MAY-2000 (first entry)		
DE	Human MDR-1 DNA.		
XX	Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;		
XX	stomach; colon; ulcerative colitis; pouchitis; large intestine; human;		
KW	inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;		
XX	ileum; anti-ulcer; anti-inflammatory; cytostatic; MDR-1; ss.		
OS	Homo sapiens.		
PN	WO200007577-A2.		
XX	17-FEB-2000.		
PD	06-AUG-1999; 99WO-DE02426.		
XX	06-AUG-1998; 98DE-1035526.		
PR	15-DEC-1998; 98DE-1057750.		
XX	(STRE/) STREMMEL W.		
PI	Stremmel W;		
XX	WPI; 2000-195439/17.		
DR	Using phosphatidylcholine having mucous membrane protective activity to		
PT	prevent or treat large intestinal diseases, such as ulcerative colitis		
PT	or pouchitis -		
PS	Disclosure: Page 17-22; 22pp; German.		
XX	This invention describes a novel medicament containing a therapeutically		
CC	effective amount of phosphatidylcholine to treat diseases, where		
CC	phosphatidylcholine has an advantageous mucous membrane protective		
CC	activity in the large intestine. The invention describes (1) a method		
CC	to regulate presence or absence of MDR(multidrug resistance)-3		
CC	transcription in a sample from patients with suspicion of stomach and		
CC	colon diseases, where treatment with phosphatidylcholine is indicated;		



(2) a diagnostic test, to detect ulcerative colitis, pouchitis, large intestinal inflammation, Crohn's disease, diverse colitis, infections enteritis/colitis, inflammation through x-ray treatment, antibiotics, chemotherapeutics, drugs or chemicals or large intestinal carcinoma, contains MDR3-specific primers to detect mutations in the MDR3 gene family and their interactions partners (in particular transcription factors); and (3) a method to manufacture a medicament for local gene therapy in the ileum with the ideal MDR3-analogous genes for patients with missing or decreased expression of MDR3-analogous proteins. The products of the invention have anti-ulcer, anti-inflammatory and cytoprotic activity. The orally administered delayed-release form of phosphatidylcholine prevents premature resorption and provides for targeted release in the lower section of the small or large intestine. This sequence encodes the human MDR-1 protein described in the method of the invention.

Sequence 3988 BP; 1173 A; 763 C; 988 G; 1064 T; 0 other;

Query Match	27.28;	Score 319.4;	DB 21;	Length 3988;
Best Local Similarity	69.48;	Pred. No. 3.1e-81;		
Matches 434; Conservative	0;	Mismatches 191;	Indels 0;	Gaps 0

OY	166	attattgataagaanaaccagatagataagtaacttcttcacagccgagatataaactgtaac	225
OY	166		
Db	1168	ataatgtataataaagccaagatatttgaaagctatctgaagagtgtggacaaacagatat	1227
OY	226	ataaagaagactctgtggaatttaanaatgtcttcttcaatattcatcacaagaacatctac	285
Db	1228		
OY	286	aagattctgaaaggtcctggaatccagaataaagttccgtgagaagacagtcgtctgtgt	345
Db	1288	aagatctgaaaggtctcggaaacctgtgaagctgcaagctggtgcaagacggtgtgctgtgtga	1347
OY	346	ctcaatgtgcaatggagaagatctagctagcttcacagctctcgaaggtatagatccgag	405
Db	1348	aacagttgctctgtggaagagacacacacgctccagctgtatgcaagagctctatgacccaca	1407
OY	406	gatggtcttatacatgtgtgtgaatgaatgacatagaagctttaatgtgtcggcatatcga	465
Db	1408	gagggagtgctcagtgctgtttgatgtgacagagatatattagaacataatgtaaagtcttctacgg	1467
OY	466	gacctattgtgagtggttagtctcaagagccctgttttcttgggagaccacccatctgaataat	525
Db	1468	gaatcatcttggtgtgtgtgagtcgaagaaacctgtattgtttgcacacacgctatgtaaaac	1527
OY	526	atcaagattgacaagatgtatgtatcagatgaaagagatgtagagagacaaagaaagaca	585
Db	1528	attcgctatgtgcgttgaaaatgtcatcaccatgtatgagatgtgagaagctgttcaaggaagcc	1587
OY	586	aatggtatgatttcatcatctgagagttctcctaataaattaatcatcaltgttagtggaanaa	645
Db	1588	aatggtcttgccttatcatcatgaactcctccataaattttacacccctgtgttggaagaga	1647
OY	646	ggaggtctcaaatgagttgtagggcagaanaacagagatctgcaattgtctgtgcttaattcga	705
Db	1648	gggggcacagtgtagtvggggcagaagacaaagagatccgcatctgcaaggtgcctgtgttcgc	1707
OY	706	aaccccaagattctgatttttgatgtagagctacgtctgtccctgtgattcagaagaagatga	765
Db	1708	aaccccaagatccctcgtctgtatgtaggcagtcgacgcttggacacagaagaaggaaga	1766
OY	766	gctgttcaagctgcacgtggaagaag	790
Db	1768	gtggttcaggtgctctgtgataaag	1792

RESULT	7
AA052726	
ID	AA052726 standard; CDNA; 4669 BP
XX	
AC	AA052726;

XX 24-JUN-1994 (first entry)  
DT  
XX Sequence of human multi-drug resistant gene mdr1.  
DE  
XX p-glycoprotein; multi-drug resistant gene; mdr1; ss  
KW

**Homo sapiens.**

Key	Location/Qualifiers
FH	

FT	/*tag= a
ET	

AA  
PN  
W09324613-A.

XX  
PD 09-DEC-1993

XX 14-MAY-1993: 93WO-US04707.  
PF

XX 22-MAY-1992: 92US-0887712.  
PR

XX  
BA  
(GENE-) GENETIC THERAPY INC

PA (OSSH ) US DEPT HEALTH & HUMAN SERVICES  
XX

PI Mcdonagh KT, Nlenhuis A, Tolstoshev P, VV

DR WPI; 1993-405805/50.

**XX**

PT encodes P-glycoprotein

XX

XX  
XX

CC application no. W087/0943m wherein the first 282 bp of the 5' UTR

removed. The codon YTN is unclear in the spec., and encodes the AA

CC  
XX

CC  
XX

Sequence 4669 BP; 1394 A; 893 C; 1126 G; 1254 T; 2 other;

Query Match	27.28;	Score 319.4;	DB 14;	Length 4669;
Best Local Similarity	69.48;	Pred. No. 3.4e-81;		
Matches 434; Conservative	0;	Mismatches 191;	Indels 0;	Gaps 0;

QY	166	attatttgataagaaccagatatagataaacttccacagctgatatataaccctgaatcc	225
Db	1526	ataattgataataaagcccaagatattgatacagctattcgaaaggtggtgcacaaacagatat	1585
QY	226	atagaaggaactctgagaaatttaaaatgcttcttcaattatccacgaaccatctatc	285
Db	1586	attaaggggaactcttggaattcagaatcagatcgaacttcaagttaccatctcgaaagaagt	1644
QY	286	aagatctcgaaaggtcctgaactctcagaattaagttctgagagacagctgccttggtcgt	345
Db	1646	aagatcttggaaggtgcctgaactcctgaaggtgcagaggtgcagagacggttcgcctgtgtga	1705
QY	346	ctcaatgycagtgaggagagatcagtgatccagctcttcagaggtgatatatgattccgcat	405
Db	1706	aacagtggtctgtggaagaagacacaacagctccagctgtgtcagaggtctctatgaccacca	1765
QY	406	gattgcttatcatctgttgatcgagaaatcgaaactcgaagctttaaatctgctgcattatcga	465
Db	1766	gaggggatactgatacgttctgtagagaaagataatttagaacataatgtaagtttctaag	1822
QY	466	gacacattggaagtgttttagtcaagagagcttcttgctgcgagccacatcagtaaacat	525
Db	1826	gaatcatctggtgtgtgagtcacgaagaaactgtattgttttcaccaacagatagctgaaac	1885
QY	526	atcaagatcagacgaatgatgtgactatgtatgaagaagatctgagagacgaacgaaggaa	585

Db 1886 atctgatgcgcgtgaataatgacacatgatgagatgagaagctgcaaggaagcc 1945  
Qy 586 aatgcgtatgatttcatatgaggttctctaataatcatatcatgtgagggaaata 645  
Db 1946 aatgcgtatgatttcatatgaggttctctaataatcatatcatgtgagggaaata 2005  
Qy 646 ggaagcctaattgagtgagggagcaagaacagatgcgaattgctcgtgcttgaatcga 705  
Db 2006 gggggcccaattgagtgagggagcaagaacagatgcgaattgctcgtgcttgaatcga 2065  
Qy 706 aaccacaagattcgtatcttagatgaggtcgtcgtccctggtatcagaagaagtcga 765  
Db 2066 aaccacaagattcgtcgtcgtgagggagcaagctcgtgagacacagaagaagcgaagca 2125  
Qy 766 gctgttcaagctgcaatgagaag 790  
Db 2126 gtgttcaagtgctcgtgataag 2150

## RESULT 8

AAZ49332

ID AAZ49332 standard; cDNA; 3860 BP.

XX AAZ49332;

XX 14-MAR-2000 (first entry)

DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.

XX Multidrug resistance; MDR-1; P-glycoprotein;

KM transmembrane efflux pump; haematopoietic stem cell; transduction;

KM bone marrow transplantation; chemotherapy; radiation therapy; cancer;

KM Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;

XX cytokine; wild-type; ds.

XX Homo sapiens.

OS

FH Key

FT CDS

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC

Claim 10; Page 68-70; 113pp; English.

This sequence represents cDNA encoding human wild-type multidrug resistance protein MDR-1. MDR-1 is a transmembrane efflux pump, responsible for the export of drugs from cells, particularly cancer cells. Wild-type MDR-1 shows increased

CC resistance to etoposide and decreased resistance to vinca  
CC alkaloids compared with a mutant form (AAV58187) where the Gly at  
CC position 185 is replaced by Val. The invention relates to transducing  
CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified haematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC haematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from haematopoietic stem cells, e.g., thalassemia,  
CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in haematopoietic stem cells.  
CC Haematopoietic stem cells that express MDR-1 will be protected against  
CC chemotherapeutic agents, so can be engrafted while the patient is  
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
CC provides sufficient cells to permit standard biochemical analysis.  
CC Overexpression of MDR-1 allows cytokine-driven expansion of  
CC haematopoietic stem cells by at least 10-fold compared with a maximum  
CC of 4-fold in known procedures.

XX Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

SQ

Query Match

27.0%; Score 317.8; DB 21; Length 3860;

Best Local Similarity 69.3%; Pred. No. 8, 8e-81;

Matches 433; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

Qy 166 attatgataagaaccagatagataacttccacagctgataataaactgaatcc 225

Db 1102 ataatgataagaaccagatagataacttccacagctgataataaactgaatcc 1161

Qy 226 atgaagaagacttggaatttaaaatgttcttaataatcatatcatatgaacatc 285

Db 1162 atgaagaagacttggaatttaaaatgttcttaataatcatatcatatgaacatc 1221

Qy 286 aagattcgaagctcgaatccagaatcgaatcgaatcgaatcgaatcgaatcga 345

Db 1222 aagattcgaagctcgaatccagaatcgaatcgaatcgaatcgaatcgaatcga 1281

Qy 346 ctcaatgagcagtggaagagtagtcagcttcagcttcagaggtatcatatgacg 405

Db 1282 aacagctgcttggaagagtagtcagcttcagcttcagaggtatcatatgacg 1341

Qy 406 gatgagcttcatcgtggttggaatgacatcgaatcgaatcgaatcgaatcga 465

Db 1342 gatgagcttcatcgtggttggaatgacatcgaatcgaatcgaatcgaatcga 1401

Qy 466 gacatatgagctggttggaatcgaatcgaatcgaatcgaatcgaatcgaatcga 525

Db 1402 gacatatgagctggttggaatcgaatcgaatcgaatcgaatcgaatcgaatcga 1461

Qy 526 atcaatgagcagatgagtagtcagcttcagcttcagaggtatcatatgacg 585

Db 1462 atcaatgagcagatgagtagtcagcttcagcttcagaggtatcatatgacg 1521

Qy 586 aatgcgtatgatttcatatgaggttctctaataatcatatcatatgagggaaata 645

Db 1522 aatgcgtatgatttcatatgaggttctctaataatcatatcatatgagggaaata 1581

Qy 646 ggaagcctaattgagtgagggagcaagaacagatgcgaattgctcgttgaatcga 705

Db 1582 ggaagcctaattgagtgagggagcaagaacagatgcgaattgctcgttgaatcga 1641

Qy 706 aaccacaagattcgtatcttagatgaggtcgtcgtccctggtatcagaagaagtcga 765

Db 1642 aaccacaagattcgtcgtcgtgagggagcaagctcgtgagacacagaagaagcgaagca 1701

Qy 766 gctgttcaagctgcaatgagaag 790

Db 1702 gctgttcaagctgcaatgagaag 1726

RESULT 9

AA249333  
ID AA249333 standard; cDNA: 3860 BP.  
XX  
AC AA249333;  
XX  
DT 14-MAR-2000 (first entry)  
XX  
DE Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.  
XX  
KW Multidrug resistance; MDR-1; P-glycoprotein;  
KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
KW gene therapy; gene replacement; genetic defect; thalassemia;  
KW Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;  
KW cytokine; mutant; ds.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 1.3843  
FT CDS /tag=a  
FT /product="Human G185V mutant MDR-1 protein"  
XX  
XX MO9961589-A2.  
XX PD 02-DEC-1999.  
XX PE 27-MAY-1999; 99WO-US11825.  
XX PR 28-MAY-1998; 98US-0086988.  
XX PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX PI Sorrentino B, Bunting K;  
XX DR WPI: 2000-072615/06.  
XX DR P-PSDB: AAY58187.  
XX  
XX Ex vivo expansion of hematopoietic stem cells transduced with a  
PT sequence encoding human multidrug resistance-1, used for bone marrow  
PT transplantation -  
XX  
XX Example 1; Page 79-82; 113pp; English.  
XX  
CC This sequence represents cDNA encoding human G185V mutant multidrug  
CC resistance protein MDR-1, where the Gly residue at position 185  
CC of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a  
CC transmembrane efflux pump, responsible for the export of drugs from  
CC cells, particularly cancer cells. The wild-type MDR-1 shows increased  
CC resistance to etoposide and decreased resistance to vinca alkaloids  
CC compared with the G185V mutant. The invention relates to transducing  
CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified haematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC haematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from haematopoietic stem cells, e.g., thalassemia,  
CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in haematopoietic stem cells.  
CC Haematopoietic stem cells that express MDR-1 will be protected against  
CC chemotherapeutic agents, so can be engrafted while the patient is  
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
CC provides sufficient cells to permit standard biochemical analysis.  
CC Overexpression of MDR-1 allows cytokine-driven expansion of  
CC haematopoietic stem cells by at least 10-fold compared with a maximum  
CC of 4-fold in known procedures.  
XX  
XX Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;  
SQ

Query Match 27.0%; Score 317.8; DB 21; Length 3860;

Best Local Similarity 69.3%; Pred. No. 8.8e-81;  
Matches 433; Conservative 0; Mismatches 192; Indels 0; Gaps 0;  
QY 166 attattgtaagaaccacagatagataactttccacagctgcatatataaccctgaatcc 225  
DB 1102 ataattgataaagacagatttgcagctatttgcagaagtgggcaacaaccagataat 1161  
QY 226 atagaaggaactgtggaatttaaaatgttcttcaatttcctacacagccatctatc 285  
DB 1162 attaagggaataatttgaattcagaatattgttcacttcagatccctccgaaagaagt 1221  
QY 286 aagatttcgaaggctcgtcaatccagaaatgaagtcgtgagagacagctgcgtgctgt 345  
DB 1222 aagatttcgaaggcctgaacctgaagtgagagtggtggtgagagagcgtggtccctgttga 1281  
QY 346 ctcaatgtcagtggaagaagtaagtaagtcagcttcctgcagaggttatatgaccggat 405  
DB 1282 aacagtggtgtggaagaagacacacagctcagctgtagcagaagctctatgacccaca 1341  
QY 406 gatgcttatcatgtgtgtagagagatgacatcagagcttaaatgtgcgcatatcga 465  
DB 1342 gaggggatggtcagtgtagagagacagatatagacacataaagttaagtttctacg 1401  
QY 466 gaccatattggaagtgtgtagtaagagcctgtttgttcggagaccacatagtaaacat 525  
DB 1402 gaatcattgtgtgtgtgtagtaagagacactgtattgtttgcaccacagatagctgaaac 1461  
QY 526 atcaagtatgagagagatgagtagtctatgaagaagatggagagagacagaggaaca 585  
DB 1462 attcgtatggtcgtggaatgtcaccatgtgtagtgaagaagcttcaagaagcgc 1521  
QY 586 aatgctatgattatatacatggaattctcctaataattatcatatgttaggggaaaaa 645  
DB 1522 aatgctatgacttatcatatgaactgcctcctaataatttgcacacctgtgtggagaga 1581  
QY 646 ggaagctcaatgagtggagggcagaacaagagatgcgaattgtcgtgcttaagtcga 705  
DB 1582 ggggccagtgtagtggggcagaagagatgcgcatgtgcagctgacctgttcgc 1641  
QY 706 aacccaagattctgattttagatgagctacgctgcctcgtgattcagaagacagca 765  
DB 1642 aacccaagatcctcctctgtgatlgagccacgctcagccttggacacagaagaagca 1701  
QY 766 gctgtcaagctgacatgagaag 790  
DB 1702 gtgttcaggtgtgctctgtgataag 1726  
RESULT 10  
ID AAV6533  
ID AAV6533 standard; cDNA: 4264 BP.  
XX  
AC AAV6533;  
XX  
DT 08-JAN-1999 (first entry)  
XX  
DE Mutated human P-glycoprotein gene designated mdr-delta-F335.  
XX  
KW mdr1-delta-F335; cyclosporin-resistant deletion mutant; human;  
KW P-glycoprotein; cyclosporin; drug efflux;  
KW modulator-resistant multidrug resistance; cancer therapy; ss.  
OS Homo sapiens.  
XX  
XX US5830697-A.  
XX PN 03-NOV-1998.  
XX PD 21-JAN-1997; 97US-0784649.  
XX PF 21-JAN-1997; 97US-0784649.  
XX PR 21-JAN-1997; 97US-0784649.  
XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

```
XX PI Chen G, SIKic BI;
XX DR WPI: 1998-609244/51.
XX PT DNA encoding mutant P-glycoprotein polypeptide - useful for
XX PT producing recombinant polypeptide
XX PS Disclosure: Columns 15-20; 17pp: English.
XX CC The present sequence represents mdrl-delta-F335, a gene encoding a
XX CC cyclosporin-resistant deletion mutant of human P-glycoprotein lacking
XX CC Phe335. The recombinant protein can be used to produce antibodies. It
XX CC can also be used in drug screening assays to study the effects of
XX CC modulators such as cyclosporin on P-glycoprotein-mediated drug efflux.
XX CC The nucleic acid can be used to transfer modulator-resistant multidrug
XX CC resistance into transfected cells, e.g. haematopoietic stem cells for
XX CC use during cancer therapy.
SQ Sequence 4264 BP; 1263 A; 807 C; 1041 G; 1153 T; 0 other;

Query Match 27.0%; Score 317.8; DB 19; Length 4264;
Best Local Similarity 69.3%; Pred. No. 9, 2e-81;
Matches 433; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 166 attatgataagaaccgcatagataactttccacagctgagatataaactgatacc 225
DB 1240 ataattgataataagccaagltatgacagctatcgaagatgaggacaacaacagataat 1299
QY 226 atagaaggaactgctggaatttaaaaatgtctcttcattatcatcacaagaaccatcat 285
DB 1300 attaagggaatttggaattcgaatcgaatgttcaacttcagttaccacactcogaanaaglt 1359
QY 286 aagatctgaaaggtcgtgaatcagaatlaagtcgtgagagacagtcgcctgtgtcgtg 345
DB 1360 aagatcttgaaggcctggaacctgaaggtgcagagtgaggcagcggtgcccgtgtgga 1419
QY 346 ctcaatggaagtgaggagagctacgtaagtcagcttcgagaggttatatagtcggat 405
DB 1420 aaaaagtgctggtggagagagcacaacagtcacgtcgaatgcagaaggtctatgacccaca 1479
QY 406 gatggtcttatactggtgagatgaaatgacatcagagctttaaattgctgacattatcga 465
DB 1480 gagggagatgtgtcagtggtgacagcagatataagaccataaattgaagttcttcag 1539
QY 466 gaccatattgaggtgtgtgacagaagcctgtttgttcgggacacacatcagtaacaat 525
DB 1540 gaatcatctgtgtgtgtgacagaaacctgtatgtttgccacacagatagctgaaac 1599
QY 526 atcaagtaaggacgaatgataatgactgataagagaatgagagacagaagaaggagca 585
DB 1600 atcgcgtatggtgcgtgaaaatgacacatgagatgagatgagaagctgtlcaaggagac 1659
QY 586 aatgcgtatgatttatactatgaggttcttctaataatlaataatctgtgtaggggaaaaa 645
DB 1660 aatgcctatgacttatactatgaactgcctcctaataattgacacccctgttggagagaa 1719
QY 646 ggaagctcaaatgagtgaggagcagaacagaagatgcgaattgctgtccttaagtcca 705
DB 1720 gggggccaagtgtggtgtgagcagaagcagatgcacattgcacgtgcccgtgttcc 1779
QY 706 aaccccaagaattctgatttagatgaggtcagtcgcctcctgattcagaagaagtcga 765
DB 1780 aaccccaagaatcctcctgtctgattgagtcacacgtcagccttgacacagaagaagcga 1839
QY 766 gctgtcaagctgcaactgaggaag 790
DB 1840 gtggttcaagtggtcctctgataag 1864

RESULT 11
AAV66534
```

```
ID AAV66534 standard; cDNA; 4264 BP.
XX AC AAV66534;
XX AC 08-JAN-1999 (first entry)
XX DT 08-JAN-1999 (first entry)
XX DE Mutated human P-glycoprotein gene designated mdr-delta-F336.
XX KW mdrl-delta-F336; cyclosporin-resistant deletion mutant; human;
XX KW P-glycoprotein; cyclosporin; drug efflux;
XX KW modulator-resistant multidrug resistance; cancer therapy; ss.
XX OS Homo sapiens.
XX PN US5830697-A.
XX PD 03-NOV-1998.
XX PF 21-JAN-1997; 97US-0784649.
XX PR 21-JAN-1997; 97US-0784649.
XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX PI Chen G, SIKic BI;
XX PI WPI: 1998-609244/51.
XX PT DNA encoding mutant P-glycoprotein polypeptide - useful for
XX PT producing recombinant polypeptide
XX PS Disclosure: Columns 27-32; 17pp: English.
XX CC The present sequence represents mdrl-delta-F336, a gene encoding a
XX CC cyclosporin-resistant deletion mutant of human P-glycoprotein lacking
XX CC Phe336. The recombinant protein can be used to produce antibodies. It
XX CC can also be used in drug screening assays to study the effects of
XX CC modulators such as cyclosporin on P-glycoprotein-mediated drug efflux.
XX CC The nucleic acid can be used to transfer modulator-resistant multidrug
XX CC resistance into transfected cells, e.g. haematopoietic stem cells for
XX CC use during cancer therapy.
SQ Sequence 4264 BP; 1263 A; 808 C; 1041 G; 1152 T; 0 other;

Query Match 27.0%; Score 317.8; DB 19; Length 4264;
Best Local Similarity 69.3%; Pred. No. 9, 2e-81;
Matches 433; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 166 attatgataagaaccgcatagataactttccacagctgagatataaactgatacc 225
DB 1240 ataattgataataagccaagltatgacagctatcgaagatgaggacaacaacagataat 1299
QY 226 atagaaggaactgctggaatttaaaaatgtctcttcattatcatcacaagaaccatcat 285
DB 1300 attaagggaatttggaattcgaatcgaatgttcaacttcagttaccacactcogaanaaglt 1359
QY 286 aagatctgaaaggtcgtgaatcagaatlaagtcgtgagagacagtcgcctgtgtcgtg 345
DB 1360 aagatcttgaaggcctggaacctgaaggtgcagagtgaggcagcggtgcccgtgtgga 1419
QY 346 ctcaatggaagtgaggagagctacgtaagtcagcttcgagaggttatatagtcggat 405
DB 1420 aaaaagtgctggtggagagagcacaacagtcacgtcgaatgcagaaggtctatgacccaca 1479
QY 406 gatggtcttatactggtgagatgaaatgacatcagagctttaaattgctgacattatcga 465
DB 1480 gagggagatgtgtcagtggtgacagcagatataagaccataaattgaagttcttcag 1539
QY 466 gaccatattgaggtgtgtgacagaagcctgtttgttcgggacacacatcagtaacaat 525
DB 1540 gaatcatctgtgtgtgtgacagaaacctgtatgtttgccacacagatagctgaaac 1599
```



XX 12-APR-1990.  
 PD  
 XX 05-OCT-1988; 880P-0251475.  
 PF  
 XX 05-OCT-1988; 880O-0251475.  
 PR  
 XX (SUNR ) SUNTORY LTD.  
 PA  
 XX WPI: 1990-159707/21.  
 DR P-PSDB; AAR04868.  
 XX  
 XX MDR related gene derived from human normal cells -  
 PT codes specific amino acid sequence, used for diagnosing drug  
 PT resistance of cancer cells  
 PT  
 PS Disclosure; ; Japanese.  
 PS  
 XX  
 CC The gene is useful for diagnosis of drug resistance of cancer  
 CC cells.  
 CC  
 XX  
 SQ Sequence 4378 BP; 1313 A; 819 C; 1058 G; 1188 T; 0 other;

Query Match 27.0%; Score 317.8; DB 11; Length 4378;  
 Best Local Similarity 69.3%; Pred. No. 9,4e-81;  
 Matches 433; Conservative 0; Mismatches 192; Indels 0; Gaps 0;  
 OY 166 attatgataagaaccagatagataaactttccacagctgataataactgatacc 225  
 DB 1239 ataattgataataagcaagatattgacagctatcgaagagtgycacaaaccagataat 1298  
 OY 226 atagaagaacgtggaatttaaaatgttcttcaatcatcaagaacatctatc 285  
 DB 1299 attaagggaattggaattcagaatgttcaacttcaagttaccacatcgaagaagatt 1358  
 OY 286 aagattctgaagaagtcagatcagaatgaattgaagtgagaagcagctgctgctg 345  
 DB 1359 aggaattctgaaggctcgaaccggaagtgcaagtgaggaagcagctgctgctgga 1418  
 OY 346 ctcaatgagcagtggaagagtgacgtagtccagctctcgaaggtatatagtatccgcat 405  
 DB 1419 aacagtgctgtggaagagcacaacagctcagctgtagcagaggtctatgagccca 1478  
 OY 406 gatggtctatcatgtgtgataagatgacatcagaagccttaattgctgcatatcga 465  
 DB 1479 gaaggagtggt 1538  
 OY 466 gacacattggaagtggt 525  
 DB 1539 gaataatcgt 1598  
 OY 526 atcaaatgagagagatgtagtcatgataagagatgagaagagcagaagaaggaagca 585  
 DB 1599 attcgtgtatgagcgtggaataatgacacatgagatgagatgagaagcgtgcaaggaagcc 1658  
 OY 586 aatgctatgatttcatcatgaggttctctaataatlaatacatatgtagaggaagaa 645  
 DB 1659 aatgctatgatttcatcatgaggttctctaataatlaatacatatgtagaggaagaa 718  
 OY 646 gagagctcaaatgagtgagagcagaagaagatgcaattcctgtgcttaattgga 705  
 DB 1719 gagagcccaagtgagtgagtgagcagaagaagatgcaattcctgtgcttaattgga 1778  
 OY 706 aaccacaagattcgtatttagatgagctagctgctcctgattcagaagaagcaagta 765  
 DB 1779 aaccacaagattcctcctgctgagtgagcagctgagccttgcagacagaagaagcaagca 1838  
 OY 766 gctgttcaagctgacactgagaag 790  
 DB 1839 gtgtgtcagtggtgctctgataag 1863

RESULT 14  
 AA072872  
 ID AA072872 standard; cDNA; 4646 BP.  
 XX  
 AC AA072872;  
 XX  
 DT 21-JUN-1995 (first entry)  
 XX  
 DE Human multidrug resistance gene (MRD-1) encoding a p-glycoprotein.  
 XX  
 XX Human; tyrosinase; compound; hair; follicle; liposome; lipophilic;  
 KW lipophilic; stratum corneum; melanin; hair dye; glycoprotein; alopecia;  
 KW chemotherapy; transforming growth factor; growth stimulant; aromatase;  
 KW cyclosporin A; macromolecule; polymer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 425..4267  
 FT /tag= a  
 FT /product= p-glycoprotein  
 FT  
 PN W09422468-A.  
 XX  
 PD 13-OCT-1994.  
 XX  
 PF 01-APR-1994; 94MO-US03634.  
 XX  
 PR 02-APR-1993; 93US-0041553.  
 PR 13-JAN-1994; 94US-0181471.  
 XX  
 PA (ANTI-) ANTICANCER INC.  
 XX  
 PI Li L, Lishko VK;  
 XX  
 DR WPI: 1994-332816/41.  
 DR P-PSDB; AA072872.  
 XX  
 PT Liposomes for delivering protein, nucleic acid etc. to hair  
 PT follicles - e.g. to restore hair colour, prevent hair loss during  
 PT chemotherapy, stimulate hair growth etc.  
 XX  
 PS Claim 8; Page 70-77; 100pp; English.

CC The nucleotide sequence of the human multidrug resistance (MDR-1) gene  
 CC which encodes a p-glycoprotein. This is an example of a compound which  
 CC can be delivered to hair follicles via a novel liposome composition.  
 CC The liposomes are comprised of a lipophilic or lipophobic compound which  
 CC will selectively target the hair follicle (via the stratum corneum)  
 CC without damaging or unwanted effects on cells outside the follicle.  
 CC Compounds e.g. p-glycoprotein, can be delivered to the hair follicle for  
 CC treatment of chemotherapy-induced alopecia. Other compounds targeted at  
 CC hair follicles can include: tyrosinase (or the DNA encoding it -  
 CC AA072872), melanin or hair dyes (to restore hair colour or condition);  
 CC human transforming growth factor-alpha (AA072873) (for reversal of wavy  
 CC hair); hair growth stimulants (e.g. cyclosporin A or aromatase) or  
 CC antisense sequences. The method allows compounds (e.g. macromolecules  
 CC or polymers), which would not normally reach the hair follicles, to be  
 CC delivered to these target areas.

SO Sequence 4646 BP; 1371 A; 893 C; 1128 G; 1254 T; 0 other;

Query Match 27.0%; Score 317.8; DB 15; Length 4646;  
 Best Local Similarity 69.3%; Pred. No. 9,6e-81;  
 Matches 433; Conservative 0; Mismatches 192; Indels 0; Gaps 0;  
 OY 166 attatgataagaaccagatagataaactttccacagctgataataaactgatacc 225  
 DB 1526 ataattgataataagcaagatattgacagctatcgaagagtgagcacaacagataat 1585  
 OY 226 atagaagaacgtggaatttaaaatgttcttcaatcatcaagaacatctatc 285







**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:46:00 ; Search time 14771.4 seconds  
(without alignments)  
1312.277 Million cell updates/sec

Title: US-09-873-409-11  
1175  
Perfect score: 1 cctaattccctaataatc.....taaaatgctgacttatgga 1175  
Sequence:

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_hig:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_higo\_hum:\*  
31: em\_higo\_inv:\*  
32: em\_higo\_rod:\*  
33: em\_hig\_hum:\*  
34: em\_hig\_inv:\*  
35: em\_hig\_rod:\*  
36: em\_hig\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	388	33.0	79611	9	AC002486	AC002486 Human BAC
2	325.4	27.7	3987	10	CRUPGP1185	M59254 Chinese ham
3	325.4	27.7	4296	10	CRUPGP1185	M60040 C.griseus P
4	325.4	27.7	4304	10	CRUPGP1165	M59253 Chinese ham
5	324.2	27.6	4045	12	AF692324	AF692324 Synthetic
6	324.2	27.6	4219	6	AX105057	AX105057 Sequence
7	324.2	27.6	4219	6	AX105078	AX105078 Sequence
8	324.2	27.6	4219	6	AX105080	AX105080 Sequence
9	324.2	27.6	4219	6	AX105082	AX105082 Sequence
10	322.6	27.5	4317	4	AF045016	AF045016 Canis fam
11	322.6	27.5	4317	6	AX105059	AX105059 Sequence
12	319.4	27.2	3988	6	AX024454	AX024454 Sequence
13	319.4	27.2	4192	6	AF016535	AF016535 Homo sapi
14	319.4	27.2	4378	6	E02326	E02326 Multidrug r
15	317.8	27.0	4264	6	AR051647	AR051647 Sequence
16	317.8	27.0	4264	6	AR051650	AR051650 Sequence
17	317.8	27.0	4646	6	I49610	I49610 Sequence 2
18	317.8	27.0	4646	9	HUMMDR1	M14758 Homo sapien
19	317.8	27.0	4669	6	AR091275	AR091275 Sequence
20	317.8	27.0	4669	6	I08557	I08557 Sequence 3
21	317.8	27.0	6505	6	AR028671	AR028671 Sequence
22	317.8	27.0	8630	6	AX012320	AX012320 Sequence
23	317.8	27.0	8630	6	AX012321	AX012321 Sequence
24	317.8	27.0	9318	6	AR028672	AR028672 Sequence
25	317.4	27.0	4280	10	CRUPGP11	M60041 C.griseus P
26	314.6	26.8	4669	6	AR055785	AR055785 Sequence
27	314.4	26.8	3682	10	MUSPGLY1A	M24417 Mouse phosp
28	314.4	26.8	4356	10	MUSMDR1A	M30697 Mouse multi
29	314.4	26.8	4924	10	MUSMDR1A	M33581 Mouse P-gly
30	311.2	26.5	4323	10	AF286167	AF286167 Rattus no
31	311.2	26.5	4927	10	AF257746	AF257746 Rattus no
32	309.8	26.4	3489	4	AB029153	AB029153 Felis cat
33	309.8	26.4	4195	6	AX108654	AX108654 Sequence
34	309.8	26.4	4195	6	AX108656	AX108656 Sequence
35	309.6	26.3	4233	6	AR123273	AR123273 Sequence
36	308	26.2	4234	10	RATMDR1	M81855 Rat mdr mdr
37	304.8	25.9	4298	10	MUSMDR	M14757 Mouse multi
38	303.4	25.8	4018	5	GGA9799	AJ009799 Gallus ga
39	298.2	25.4	4390	5	XL017608	U17608 Xenopus lae
40	293.8	25.0	3858	4	OA078609	U78609 Ovis aries
41	291.6	24.8	4084	10	MUSMDR1A	J03398 Mouse mdr g
42	287.4	24.5	3924	6	AX024455	AX024455 Sequence
43	287.4	24.3	3924	9	HUMMDR3	M23234 Human membr
44	285.2	24.3	3912	10	RATPGLYCO	L15079 Rat P-glyco
45	283.6	24.1	3905	10	CRUPGP111	M60042 C.griseus P

## ALIGNMENTS

RESULT	1	LOCUS	AC002486	DEFINITION	Human BAC clone CTA-367017 from 7p15-p21, complete sequence.	VERSION	AC002486.1	GI:2341014	ORGANISM	Homo sapiens
REFERENCE		AUTHORS	Kalicki,J., Smith,A. and Gibson,A.	TITLE	The sequence of H. sapiens BAC clone CTA-367017	JOURNAL	Unpublished	2 (bases 1 to 79611)	REFERENCE	Submitted (22-AUG-1997) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 3 (bases 1 to 79611)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 4 (bases 1 to 79611)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT  
 -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@wustl.edu](mailto:sapiens@wustl.edu)  
 -----  
 Summary Statistics  
 -----  
 Center project name: H\_RG367017  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7> or send mailto:egreen@nhgri.nih.gov

SOURCE INFORMATION:  
 This clone is from a release of the human BAC library. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).  
 VECTOR: pBelOBAC11  
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the right is CTB-86D3. The actual start of this clone is at base position 1 of CTA-367017. This clone is part of an unanchored island, orientation is unknown.

This clone contains STS SWSS1805 (NID:g1113222).

#### FEATURES

source  
 1..79611  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /clone="CTA-367017"  
 /clone\_lib="CTB-HS-A"  
 /map="7p15-p21"  
 683..715  
 /rpt\_family="L1"  
 repeat\_region  
 1781..2410  
 /rpt\_family="L1"  
 repeat\_region  
 3374..3875  
 /note="probable ras-related (RAL) pseudogene: similar to (P1D:g1346852)"  
 misc\_feature  
 3429..3529  
 /note="similar to EST T29472 (NID:g611570)"  
 misc\_feature

misc\_feature  
 3550..3804  
 /note="similar to EST AA523370 (NID:g2264082) n167d09.s1"  
 misc\_feature  
 3810..4260  
 /note="similar to EST AA403045 (NID:g2055607) zv63c09.r1"  
 misc\_feature  
 3903..4207  
 /note="similar to EST Z44835 (NID:g574005)"  
 misc\_feature  
 complement(3910..4337)  
 /note="similar to EST N62870 (NID:g1210699) yz83e03.s1"  
 misc\_feature  
 4025..4431  
 /note="similar to EST T74360 (NID:g691035) yc84h10.r1"  
 misc\_feature  
 4217..4745  
 /note="similar to EST AA449169 (NID:g2162632) zx02g12.r1"  
 misc\_feature  
 4265..4572  
 /note="similar to EST F12824 (NID:g708845)"  
 misc\_feature  
 /note="similar to EST AA382990 (NID:g2035360)"  
 4546..4861  
 /note="similar to EST AA296743 (NID:g1949076)"  
 misc\_feature  
 complement(4639..4792)  
 /note="similar to EST AA368261 (NID:g2020654)"  
 misc\_feature  
 4733..4933  
 /note="similar to EST AA447258 (NID:g2159923) zw93e03.r1"  
 misc\_feature  
 4938..5215  
 /note="similar to EST AA447258 (NID:g2159923) zw93e03.r1"  
 misc\_feature  
 complement(5063..5638)  
 /note="similar to EST AA044672 (NID:g1522874) zk73a09.s1"  
 misc\_feature  
 complement(5068..5638)  
 /note="similar to EST AA044672 (NID:g1522874) zk73a09.s1"  
 5145..5435  
 /note="similar to EST H97948 (NID:g1118833) yx04e05.s1"  
 misc\_feature  
 5147..5571  
 /note="similar to EST W56187 (NID:g1358144) zd14c10.r1"  
 repeat\_region  
 complement(5642..5673)  
 /rpt\_family="L1"  
 5782..6074  
 /rpt\_family="ALU"  
 repeat\_region  
 6499..6859  
 /rpt\_family="L1"  
 repeat\_region  
 9618..9657  
 /rpt\_family="L1"  
 repeat\_region  
 /rpt\_family="L1"  
 complement(10090..10285)  
 /rpt\_family="L1"  
 repeat\_region  
 13808..13899  
 /rpt\_family="L1"  
 repeat\_region  
 complement(16807..16844)  
 /rpt\_family="L1"  
 repeat\_region  
 complement(18057..18887)  
 /rpt\_family="L1"  
 repeat\_region  
 18866..20301  
 /rpt\_family="L1"  
 repeat\_region  
 /rpt\_family="L1"  
 complement(19766..19812)  
 /rpt\_family="L1"  
 repeat\_region  
 complement(20583..20874)  
 /rpt\_family="ALU"  
 repeat\_region  
 complement(21847..22287)  
 /rpt\_family="L1"  
 repeat\_region  
 21868..22285  
 /rpt\_family="L1"  
 repeat\_region  
 complement(22880..22983)  
 /rpt\_family="ALU"  
 repeat\_region  
 23346..23476  
 /rpt\_family="L1"  
 repeat\_region  
 23578..24277  
 /rpt\_family="L1"  
 repeat\_region  
 24960..25002  
 /rpt\_family="L1"  
 repeat\_region  
 25076..25163  
 /rpt\_family="L1"  
 repeat\_region  
 25210..25280  
 /rpt\_family="L1"  
 repeat\_region  
 25304..25362  
 /rpt\_family="L1"  
 repeat\_region  
 25420..25459  
 /rpt\_family="L1"  
 repeat\_region

```
repeat_region      /rpt_family="LI"      25489. .25965
repeat_region      /rpt_family="LI"      25968. .26249
repeat_region      /rpt_family="ALU"      26287. .26341
repeat_region      /rpt_family="ALU"      26702. .27068
repeat_region      /rpt_family="LI"      27089. .27207
repeat_region      /rpt_family="ALU"      27223. .27260
repeat_region      /rpt_family="ALU"      27261. .28126
repeat_region      /rpt_family="LI"      .30118
repeat_region      /rpt_family="ALU"      .32163
repeat_region      /rpt_family="LI"      .32565
repeat_region      /rpt_family="LI"      .32566
repeat_region      /rpt_family="ALU"      .32857
repeat_region      complement(32858. .33932)
repeat_region      /rpt_family="LI"      33937. .35195
repeat_region      /rpt_family="LI"      36419. .36448
repeat_region      /rpt_family="LI"      39372. .40123
repeat_region      /rpt_family="LI"      40278. .40297
repeat_region      /rpt_family="LI"      41686. .41708
repeat_region      /rpt_family="LI"      .42243
repeat_region      complement(41775. .42243)
```

Query Match	33.0%	Score 388	DB 9	Length 79611
Best Local Similarity	100.0%	Pred. NO. 1.1e-78		
Matches 388	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 788	aggataccccagagatcatctttagactaaatttaacctcaagtggagatcgtagact	847		
Db 72847	AGGATACCCCCAGGATTCATTGTGACCTAATTTCACCTCAAGTGGAGATCGCTGACCT	72906		
QY 848	tgaacagagcgcccttcgacagctctgagccctcaaacctacccctgaactcctgtgctc	907		
Db 72907	TGAACACACGCGCCCTTCGACAGCTCTGGGCCCCCAAAACCTCACCGACCTCCTGCGGCT	72966		
QY 908	atgagctactctgacatacctaactaaagcagatagcagtttggcgcccgaccacaaatracct	967		
Db 72967	ATGAGCTACTGTGCACATTCCTCAAGGCGATATCAGATTGTGGCGCTGGACCAATTTACCT	73026		
QY 968	gaatcaagagagggaggttggcagttggcggtatgaataaacaatlgacagtttcttcagag	1027		
Db 73027	GAATCTAGGAGGAGGAGTTGGCAGTGGCGGATGAAAAACCATTTGAACAATTTTCGGAG	73086		
QY 1028	gctgtagctcccttataaaccagagcccttcagagcccttaacaaggtcttaatggacatttt	1087		
Db 73087	GCTGTAGCTCCCTTAATAAACCAAGAGCCTTCAGACCCCTTAACAAGGCTTAATGCGACATTTT	73146		
QY 1088	acttgcgaatttgccttggagaagtgaagtcagtttcttctctcaagaanaatcgacaggt	1147		
Db 73147	ACTTTGCATTTGCTCTGGAGAGTGAAGCTTAAGCGTTTCTTCTCTAAGAAAATCGCAGGCT	73206		
QY 1148	tctttttttaaaatgctgactttatgga	1175		
Db 73207	TCTTTTAAAAATGCTGACTTTATGGA	73234		

RESULT 2

CRUPEP1185 CRUPEP1185 CRUPEP1185

CRUPEP1185 3987 bp

CRUPEP1185 mRNA

CRUPEP1185 ROD

27-APR-1993

DEFINITION	Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
ACCESSION	M59254
VERSION	M59254.1 GI:191156
KEYWORDS	multidrug resistance glycoprotein; p-glycoprotein.
SOURCE	Chinese hamster, cdna to mRNA, clone ADX185.
ORGANISM	Cricetus sp.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetus.
AUTHORS	1 (bases 1 to 3987)
TITLE	Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
JOURNAL	Full length and alternatively spliced ppv-1 transcripts in multidrug resistant Chinese hamster lung cells
MEDLINE	J. Biol. Chem. 266, 4545-4555 (1991)
FEATURES	91154265
source	Location/Qualifiers
gene	1..3987
CDS	/organism="Cricetus sp." /db_xref="taxon:10031" /cell_type="fibroblast" 115..3624 /tissue_type="lung" 115..3624 /gene="ppv-1" 115..3624 /gene="ppv-1" /product="p-glycoprotein" /codon_start=1 /protein_id="AA37005.1" /db_xref="GI:191157" /translation="MTTAAAYTGTGAGVLTAAIVQVSWFLAGROIHKTROFEPHAI IMNEIGWEDVDVDELITRLIDVYSKINEGIGKIGAFVQAAFPFGGITFTTRGC KLTLLVILSPVGLISAGIMAKLILSPFKDELQAVNAGVAAEVLAIIRVIAFGGQ KGLTEYRNNNLNEAKRKILGKAITATINSMGAFLLIAYAVLAWYGSITVMEYS GOVLEVEFVAILAEFSGIQAIPNEAFNAGFAVEIFINIDNKSIDSEFKNYKDD NIKCNLEFNKHFSTYPSKRDDVOILKGNKQSOATVALAENIRGENTVMEIEK DPTEGVSIDSGODRTINVRILERTIGVSEPELFTTVAENIRGENTVMEIEK AVKNAADITMKLPKRPDITVGERGALSGGQRIARALVRNPKILLDATSA LDITSEAVQALDKAREGRTTIVIAHLSTVRNADITAGDGVIVQGHLELMRE KGIETFLMTQAGNEIELGNVEGSKEDINLDMSSDASLSLRSTPSSIRGPRE DODRLSTREKALDDVPISEWRIKLKLSSEMPYFVGAICAYNALOPAFSIFES VVGFTNRTNDETRHDSNPLSLFLILGVTSTFPLFGCTFGKAGILTKRLRYMW FKSLRDVSWFEDNPKNTGALITRLANDAGVAGCARLAVITNANTGCTIIS LITVQKLLTLLALVPIIAIAGVEMKLSGALKDKKELGSGKITEATIENTRTV SLTPEOKFENMYAGSLQIPYRNALKAHVEFTSFQAMVFSYACFRGALVANR ELMTFENVILVESAIVGMAVGVSSSPADPYARAKASAHIMIEKVPISDSYTG GLKNTLEGANVKNFENVENYPTRPDIPVLOGILNEKKGOTLALVSGSGKSTVVO LERYDPMAGKTVPLIDGKEVNOLNOMRAHGIYSOEPLIFDCSIAMIAVGNMSHL SODETERAKNANTHOFIESLPDKYINRVGVCKGTOLSGGQRIARALVRNPKIL LDITSEALDITSEKVVQALDKAREGRTTIVIAHLSTVQNALVIAVQNGVEHGT HQQLLAQGIETFSWVSVOGAKR"
BASE COUNT	1194 a 822 c 944 g 1027 t
ORIGIN	Map position 1q26.
Query Match	27.7%; Score 325.4; DB 10; Length 3987;
Best Local Similarity	68.2%; Pred. No. 1.9e-64;
Matches	452; Conservative 0; Mismatches 211; Indels 0; Gaps 0;
Db	131 ccatgcagtatctgcatgtgagcagcagccctcaatatctgataagaaccagtatag 190
Db	851 CCAATGCAAGAGAGACACGCTTATGAATCTTAAACATATTTATATATAGCCCACTAATTG 910
Db	191 ataactttccacagcagctgatatataaccctggaatccatagagaagcctgagaaat 250
Db	911 ACAAGCTTCCAAAGATGGGTCAAAACCGACACATTTAAAGGAAATTGGAAATTCMAAA 970
QY	atgtttcttcaatatcatcaataagaaccatctatcaagaattctgaaagctgaaatc 310
Db	971 ATATTCACTTCAGTTACCCATCTCGAAAAGACGTTACAGATCTTGAAGGCGCTCAACCTGA 1030
QY	311 gaattaagctctggaagaagacgcctccttggtgctcctcaatgagcagtggaagagta 370
Db	1031 AGTGTCAGCGCGAAGACAGTGGCCCTGGTGGCAACAGTGGCTGTGGGAAGAACACCA 1090

Oy	371	tatgcacagcttcctcgcagaaggtctatacatgaccgaatgatggcttatacaatgtagtgatgaga	430
Db	1091	CTGTCCACCTGCTCGACAGGCTTATCAGACCACCCACAGAAGGGCGTGCTCATGATCGACAGGAC	1150
Oy	431	atgacatcacagactttaattgtagcgcatltaacgagacatatlttgagagtgttagtcaag	490
Db	1151	AGGACATCAGGACCATCAATGTGAAGGTATCTGCCGGAATAATTGGGGTCGTGATGACAG	1210
Oy	491	agacctgtttgttcggagccaccatcagtaaacatatcaatgatgatgagcagatgatgtga	550
Db	1211	AACCTGTGTTGTTGGCCACCACTATCCGTGAAAACATTCCGTATGAGCCGACAATAATGCA	1270
Oy	551	ctgatgaagatgatgagagagcagcagcaagaaggacaatgtgcatgatatttatcagtgagt	610
Db	1271	CTATGAGACGAGATCCAGAAAGCTGTCAAGSAGAACCCATTGCCTATGACTTCATCATGAAC	1330
Oy	611	tctcctaataatttaatacatctgtgtagaggaaaaagagctccaatgagtgtgagggcaga	670
Db	1331	TGCCCATTAATAATTGACACTCTGTGTGGTGAAGAGAGGACACAGCTGATGAGAGGACAGA	1390
Oy	671	aacagagatgcgaattgtctcgttccttagtgcgaaccccagaagattctgatttagatg	730
Db	1391	AACACAGATGCGCCATTCCTGCTGTCGCCGCTGTCGCCGACCCCAAGATCTCTTGTGTGATG	1450
Oy	731	aggctacgtctgcccttgatttcagaagaagaagctagctgttcaacgtcactgcaggaaga	790
Db	1451	AGGCCACATCAGACCTTGGAGCACACAGAAAGTGAAGCCGTTCAGGCTCTCTGATTAAG	1510
Oy	791	ata 793	
Db	1511	CYA 1513	
RESULT	3		
LOCUS	CRUGP1	4296 bp mRNA	ROD 22-JUN-1995
DEFINITION	C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.		
ACCESSION	M60040		
VERSION	M60040.1 GI:191164		
KEYWORDS	multidrug resistance; p-glycoprotein; transmembrane protein.		
SOURCE	C.griseus adult liver and ovary, cDNA.		
ORGANISM	Cricetulus griseus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;		
	Cricetulus.		
REFERENCE	1 (bases 1 to 4296)		
AUTHORS	Endicott,J.A., Sarangi,F. and Ling,Y.		
TITLE	Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene family		
JOURNAL	Data Seq. 2 (2), 89-101 (1991)		
FEATURES	92135896		
source	Location/Qualifiers		
	1..4296		
	/organism="Cricetulus griseus"		
	/db_xref="taxon:10029"		
	/dev_stage="adult"		
	/tissue_type="liver and ovary"		
	103..3933		
	/note="putative"		
	/codon_start=1		
	/product="P-glycoprotein isoform I"		
	/protein_id="AAA6883.1"		
	/db_xref="GI:191165"		
	/translation="MEPEEDSSGRKKDKNLKMGSKSEKEKRPVSYLPMFPYADQ LDRLMLGTLAIIHGVALPLMVLVEDMDTDSFASGNITNTANNATOVNASDING KLEEMTYAYVTYGIGAGVLIIVAYIOVSFWCLAAQRQIRIKRPFHALIMNODIWMF DVHGELNTRLTDDSVSKINDGIDKIQMFQOAMATEFGFIIGFTRWKLTVILAD SPVLSISAGIWAKLISLSEPKLOAVAKAGVAEEVLAARTVIAFGCKEELRYNNK NLEEKRRGIKKAITANTSMGAFLIIVASYALAIFWGETSLIVISEKISGOVLVFPF VLICAFSTIGASPINEAPANNRGAAVEIFNIIDUKPSIDSRSKNGYPRDNKJLERK NIHSYPRKDVOILKGLNVLSVGQVALVANGSCGSGSTIVQLDRITDTBEVAS GDQDIRIVARYLRRIIGVNSDEPLVATPTTIENIRGRENVTLDEKAVEKANAYDI		
CDS			

BASE COUNT	1271 a	892 c	1029 g	1104 t
ORIGIN				
Query Match	27.7%	Score 325.4	DB 10	Length 4296
Best Local Similarity	68.2%	Pred. No. 2e-64		
Matches 452	Conservative 0	Mismatches 211	Indels 0	Gaps 0
OY	131	ccatgacgattctgcatctgagcagcagctccctcatattgataagaaccagatag	190	
Db	1160	CCAATGCAGAGAGAGACCAACCCATGAAATCTTCACATATATTGATATATAGCCAGTATG	1219	
OY	191	ataactttccacagcctgagatataaacctgaatccatagaaagaaactctggaatttaa	250	
Db	1220	ACAGCTTCTCAAGAGATGGGTACAAACCCAGACAAACATTAAAGAAATTTGGATATCAAA	1279	
OY	251	atgttcttcaattatcatcatcaagaaccatcatcaagaattctgaaagctcgaattcca	310	
Db	1280	ATATTCACCTTCAGTTACCCATCTCGAANAAGACGTTCAATCTTGAGAGGCGCTCAACCTG	1339	
OY	311	gaattaaactcggagagagacatcgcccttgctgctcgaatgagcagctgggaagtagc	370	
Db	1340	AGGTGCAGAGAGGAGACAGACAGTGGCGCTTGTCGCAACAGTGGCTGTGGAAAGCACCA	1399	
OY	371	tagtcacagcttcgagaggttatgatgcgagatgagcttcatcaltggtgatgaga	430	
Db	1400	CTGTCCACCTGCTGCAGAGGCTCTACGACCCACAGAGGGCGGTGCTCAGTATCGACGGT	1459	
OY	431	atgacataagactttaaltgtagcattatcgagacacatatggagtggtatgtaag	490	
Db	1460	AGGACATCAGAGCAATCAATGTGAGGTATCTCGGGGAAATCTTTGGGCTCGTAGTCAAG	1519	
OY	491	agccctgtttgctcgagacacacacatcaatcaatcaatgaatgatgagcagagatgta	550	
Db	1520	AACCTGTGTGTTGCCACACACTATTCGCTGAANAACATTCGCTATGCGCGAGAAATGTCA	1579	
OY	551	ctgagtgaagagatggagagagcagcagaaggaacaaatgtagatattatcatgaggt	610	
Db	1580	CTATGACAGCATCGAGAAAGCTGTCAAGGAGAGCCAAATGCTATATCTTCATCATGAAAC	1639	
OY	611	ttcctataaatattatcatattgtgtaggggaaagaggtcctaaltggtggagggcaga	670	
Db	1640	TGCCCATTAATTTACACTCTCTTTGGTAGAGAGAGACACAGCTGTAGTGGAGAGACGA	1699	
OY	671	aacagagatcgcaatgctcgctgcttaagtctgaaaccccaagaattcgaatttcaag	730	
Db	1700	AACACAGATTCGCCATTGCTCTGCTGCCCTGGTCCGCAACCCCAAGATCCTTTGTTGAGG	1759	
OY	731	aggtactgctgccttgagatctcagaagaacgaatcagctgtgtcaagctgcaactgsgaag	790	
Db	1760	AGGCGACATACACCTTGGACACAGAAAGTGAAGCCGTGTGATCAGGCTGCTCGATTAAG	1819	
OY	791	ata 793		
Db	1820	CTA 1822		

RESULT 4

CROP6P1165

LOCUS CRUPGp1165 4304 bp mRNA ROD 27-APR-1993  
DEFINITION Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.  
ACCESSION M59253.1  
VERSION M59253.1 GI:191154  
KEYWORDS multidrug resistance glycoprotein; p-glycoprotein.  
SOURCE Chinese hamster, CDNA to mRNA, clone ADX165.  
ORGANISM Cricetus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Cricetus.  
REFERENCE 1 (bases 1 to 4304)  
AUTHORS Devine, S.E., Hussain, A., Davide, J.P. and Melera, P.W.  
TITLE Full length and alternatively spliced pgp-1 transcripts in  
multidrug resistant Chinese hamster lung cells  
JOURNAL J. Biol. Chem. 266, 4545-4555 (1991)  
MEDLINE 91154265  
FEATURES  
source Location/Qualifiers  
1..4304  
/organism="Cricetus sp."  
/db\_xref="taxon:10031"  
/cell\_type="fibroblast"  
/tissue\_type="lung"  
111..3941  
/gene="pgp-1"  
111..3941  
/gene="pgp-1"  
/product="p-glycoprotein"  
/protein\_id="AA037004.1"  
/db\_xref="GI:191155"  
/translation="MEPEEDFGSRKDKNFKMKRKKKKRPVSVPTMERYAG  
LDRLVMTGTAATIHGVALPLMLVPGMDSDSPSNINPNNATNNATQVNASIRP  
KLEBEMTYATYYTGAGVLLVATVYSENCLAGROIHKIQFHHAMQETGMF  
DVHDGELNTRLTVDVSKINEGIGDKIGFQAMTFEGGIFGTGKMLTLVLA  
SPVLDSAGIMAKILSPTDEKELQAYAGAAVEVLAIRVIAFGQKELERYNN  
NLEEARLGIKKAITANISMAAFLLIYASALAMVSTGSLVISEKYSIGOVLEFEA  
VLIAFESIGASPNTEAFANRGAFLIPIIIDNKPSIDSESKNGYKPDNKGNEFK  
NIHPSYRNDVOILKGLIKVQSGQYALVNSGSGSTYQVLLQRLYPTREYVSI  
DGDIRTINVRILREITGVSOEPLVETFTIAENTRENTMEIKAVKEAAYD  
FTMKLPKFDLVEGSAOLSGGQKRIARALVRNKKILLDEATSLTDESAVY  
QALADKAREGRTTIVIAHRLSTVRNADVIAGDGVIVIEQGHHELMKEGTYELM  
TQAGNEIELENEVESKNEIDNDMSKDSLSLIRRSSTPGRPHDORLSTK  
EALDDVDPISFMRILKLNSSPMPEYVVICAYNGALOPAFSTIFSKVVGVRNT  
DDETRHDSNLSFLILYISSTTFPGCTPEKAEIILKRLRYRFSMLQDY  
SWFNPRTTGAALTRLANDAGVKGATGARAVTQINANGTIIISLYGWLTL  
LLAIVPIITAIAGVEMKMSGQALKDKKEGSKLATEIENRTVARSILREQKE  
NMYASLOIPYRNALKAHVGITFSFTQAMVFSYACFRGAVLVARBELTFENVL  
LVFSALIVGAAVGVSSPADYAKAKVSASHIIMIEKVSIDYSTGGLKPNLEG  
NKKNEVEVNPTRDIPVGLNLNLEVKGGOTLALVSGGSKSVVOLLEFEYPM  
GVFLDKEVNOLNOMLRHGLVTSORPILFDCSIANINAGDSRVSODEITERA  
KEANIHOTIESLPKRYNTRVGDKGTOLSGGQKRIARALVRNKKILLDEATSLD  
TESEKVQEDLKAREGRTTIVIAHRLSTVIANADLIVIQNGKEHGHQDLAOKG  
IYFSNVSVQAAKR"

BASE COUNT 1275 a 895 c 1031 g 1103 t  
ORIGIN Map position 1926.

Query Match 27.7%; Score 325.4; DB 10; Length 4304;  
Best Local Similarity 68.2%; Pred. No. 2e-64;  
Matches 452; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY 131 catagcagattatgcattgagcagcagtcctcatattatgaataaccagatag 190  
DB 1168 CCAATGCAAGAGAGAGACCTTATGAANTCTTCAACATTAATTAATTAAGCCAGATTG 1227

QY 191 ataactttccacagctgatalaataactgaatccatagaagaactgtgaaattcaaa 250  
DB 1228 ACAGCTTCCAAAGATGGGTCAAAACGACACACTTAAGGAATTTGGAAATTCAAA 1287

QY 251 atgttcttccaatatcatcaagcaactcatcaagaattcgaaggtctgaattcca 310  
DB 1288 AATATTCATTCACTGATACCATCTCGAAAGAGCTTCAGATCTTGAAGGCGCTCAACCTGA 1347

QY 311 gaataagctcgaagagacagctgccttgctgcgtcctaattgagcagtggaagatcag 370  
DB 1348 AGGTGCAGAGCGACAGACAGATGCGCTCTGTTGCAACAGTGGCTTGGAAAAGACACA 1407

QY 371 tagtcagcttcgcagagagattatgatcgcgatgatgattatcatcagtgatgaga 430  
DB 1408 CTGTCCAGCTGCGCAGAGAGCTCTACGACCCACAGAGGGCGCTGCATATCGACGAGC 1467

QY 431 atgacatcagagcttaaatgtgcgcattatcagacatattgagatggttagtcaag 490  
DB 1468 AGCAGATCAGAGACCATCATCATGTGAGGTATCTGCGGAAATCATGTCGGTCTGAGTCAAG 1527

QY 491 agcgtgttctgctggaccacatcagtaacaattcaagatattgagcagatgagta 550  
DB 1528 AACCTGTGTTGTTGCCACACATCTGCTGTAAGAAATTCGTTATGSCCAGAAATGTCA 1587

QY 551 ctgatgaagatctgagagagcagcaagaagaacaaatcgtatgattatcatcagat 610  
DB 1588 CTATGACAGCATCGAGAAAGCTGTCAAGAAAGCCATGCTATGACTTCATCATCATGAAC 1647

QY 611 ttccataataattatcatcattgtaggaggaagaaagagctcaaatgagtggagcaga 670  
DB 1648 TGCCCATTAATTTGACACTCTTGTGTGAGAGAGACACAGCTAGTGGAGGACAGA 1707

QY 671 aacagagagtcgaattgctgccttgctgagcagaaaccccaagatctgatttagatg 730  
DB 1708 AACAGAGATTCGCCATTCCTGCTGCTGCTGCGCAACCCCAAGATCTTTGTTGGATG 1767

QY 731 agctcagctgcgccctgagttcagaagaacagtcagctgctcagctcagctcagctgag 790  
DB 1768 AGCGCAGATCAGCCTTGAGACAGAAAGTGAAGCCGTGCTGAGCTGTGAGTAAGG 1827

QY 791 ata 793  
DB 1828 CTA 1830

RESULT 5  
LOCUS AF269224 4045 bp DNA SYN 04-JUL-2000  
DEFINITION Synthetic construct Canis familiaris his-tagged-multidrug  
resistance glycoprotein gene, complete cds.  
ACCESSION AF269224  
VERSION AF269224.1 GI:8926216  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 4045)  
AUTHORS Roulet, A.C., Puel, O., Gesta, S., Alvinerie, M. and Pineau, T.J.  
TITLE Analysis of dog MDR1 p-glycoprotein  
JOURNAL Unpublished  
AUTHORS  
JOURNAL  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de  
Tournefeuille, BP3, Toulouse 31931, France  
FEATURES  
source Location/Qualifiers  
1..4045  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="original mRNA isolated from Beagle dog colon"  
101..3964  
/codon\_start=1  
/transl\_table=11  
/product="his-tagged-multidrug resistance glycoprotein  
MDR1"  
/protein\_id="AA081747.1"  
/db\_xref="GI:8926217"  
/translation="MHHHHHHPDEGRGSAEKNFMWKKSKKKKKRPVSTFA  
MFRYSNMVLDRLVMTGTAATIHGVALPLMLVPGMDSDSPSNINPNNATNNATQVNASIRP  
KLEBEMTYATYYTGAGVLLVATVYSENCLAGROIHKIQFHHAMQETGMF  
DVHDGELNTRLTVDVSKINEGIGDKIGFQAMTFEGGIFGTGKMLTLVLA  
SPVLDSAGIMAKILSPTDEKELQAYAGAAVEVLAIRVIAFGQKELERYNN  
NLEEARLGIKKAITANISMAAFLLIYASALAMVSTGSLVISEKYSIGOVLEFEA  
VLIAFESIGASPNTEAFANRGAFLIPIIIDNKPSIDSESKNGYKPDNKGNEFK  
NIHPSYRNDVOILKGLIKVQSGQYALVNSGSGSTYQVLLQRLYPTREYVSI  
DGDIRTINVRILREITGVSOEPLVETFTIAENTRENTMEIKAVKEAAYD  
FTMKLPKFDLVEGSAOLSGGQKRIARALVRNKKILLDEATSLTDESAVY  
QALADKAREGRTTIVIAHRLSTVRNADVIAGDGVIVIEQGHHELMKEGTYELM  
TQAGNEIELENEVESKNEIDNDMSKDSLSLIRRSSTPGRPHDORLSTK  
EALDDVDPISFMRILKLNSSPMPEYVVICAYNGALOPAFSTIFSKVVGVRNT  
DDETRHDSNLSFLILYISSTTFPGCTPEKAEIILKRLRYRFSMLQDY  
SWFNPRTTGAALTRLANDAGVKGATGARAVTQINANGTIIISLYGWLTL  
LLAIVPIITAIAGVEMKMSGQALKDKKEGSKLATEIENRTVARSILREQKE  
NMYASLOIPYRNALKAHVGITFSFTQAMVFSYACFRGAVLVARBELTFENVL  
LVFSALIVGAAVGVSSPADYAKAKVSASHIIMIEKVSIDYSTGGLKPNLEG  
NKKNEVEVNPTRDIPVGLNLNLEVKGGOTLALVSGGSKSVVOLLEFEYPM  
GVFLDKEVNOLNOMLRHGLVTSORPILFDCSIANINAGDSRVSODEITERA  
KEANIHOTIESLPKRYNTRVGDKGTOLSGGQKRIARALVRNKKILLDEATSLD  
TESEKVQEDLKAREGRTTIVIAHRLSTVIANADLIVIQNGKEHGHQDLAOKG  
IYFSNVSVQAAKR"

GKMLTIVLAIISPVLSIAAIIMAKILISSTDKELLAYAKAGAAVEVLAARIVYANG  
 GOKKELERNKLNLEAKRIGIKRAITANISIAAFLIIVSIALAFYSGISVSSY  
 SIGOVLFVPSFVLSIGASPSIAEFAINAGAAFLIIVSIALAFYSGISVSSY  
 PDNKNLEFRNWHFSPSRKRYKILKGLKYOSSGOTVALVNGSGSGSTYOMLR  
 LYDPTDGMVCIODODITIVRHRRETTYVSOEPUVPAITIANINIGRENTYMDTI  
 EKAVKENAIDFIMKLPKFDITLVEERGAOLSGOKORAIARLVNPKLILDEAT  
 SALDSEAAVVOALDEARGRITIVIAHRLSTYRANADYAGEDDGVIVKGNDELM  
 KEGIYFKLVTOMRGENIELEENATGESKESPDLEMSDGSGLIKRSTRSIIE  
 POGDRKLTGKEDENVPVSEFWRLIKLNSTEMPYVVGICFALINGLOPASIIF  
 SRIIGITRDEDEPTKRONSMSEVLEVLVIGISITFIFLOGFPFGKAGITRLRY  
 MVRSMRLROVSMEDDPKNTTGATITTLANDAAOVKALISRLAVITONTANCTGI  
 ISLTGQOLTLILAIVPIAIGVNEKMLSGALDKDELBGALIAEALENFTI  
 VSLTRQKREYMAQSLQVYRNSLRKAHFGVSESITQAMMFYSYAGCFREAVYV  
 ANEFMNPQDLVLSAIVGMAVGVSSFAADYAKKVAANHIMIESKSPILDSY  
 PHGLKPTLEGNVTFENEVNYPTRPDIPIVLOGILEVKKGOTLIVGSGSGSTYV  
 OLLEREYDPLAGSVLIDGKEIKHLNVOMLAHGIYVSEDFILPECCSAENIYVGNR  
 VYSHREIMQAKENIHHFETLPEKNTYVGDGTLOSGOKORAIARLVNPKI  
 LLDEATSAIDTSEKVVQALDARREGRTIVIAHRLSTIONADLIVFQNKVKEH  
 GTHOOLLAQGIYFSMISVOAGAKR"

BASE COUNT 1171 a 850 c 986 g 1038 t  
 ORIGIN  
 Query Match 27.6%; Score 324.2; DB 12; Length 4045;  
 Best Local Similarity 69.9%; Pred. No. 3,7e-64;  
 Matches 437; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 166 attattgataaagaccagafatagataactttccacagctgagatataactcgaatcc 225  
 DB 1226 ATAAATGCAATTAACCAAGCATTTGACAGCTATTGGAAGCTGGCAATTAACCAAGCATTAAT 1285  
 QY 226 atagaaggaactgtagaatttaaaatgtcttcctcaatataccataagacatc 285  
 DB 1286 ATTAAGGGAATTTGGATTCGAAAATGTCTCACTCACTTACCCTTCGAAAAGAAATTT 1345  
 QY 286 aagattctgaaagtcyaaatcccaagaaatgaagtcgagagacagtcgccttgctggt 345  
 DB 1346 AAGATCTTAAGGGGTCTCAACCTGAAAGCTTCAGATGGGAGACAGTGGCCTGGTTGGG 1405  
 QY 346 ctcaatgcaatggaagaagtagtagctctcagctctcagaggtatataatccgagat 405  
 DB 1406 AACACTGGCTCGGGAAGACAGACGCTGACAGCTGATGACAGGCTGTATGACCCACACA 1465  
 QY 406 gatgcttatacatgtagatgagaaatgacacagcagcttaaatgctgcgcatatcga 465  
 DB 1466 GATGGCATGCTGTATGATGAGACAGACATTAAGCAATTAATTAAGGATCTTCGG 1525  
 QY 466 gaccatattgtagtgtagtcaagaagcctgtttgttcggagacacatcaagtaagaat 525  
 DB 1526 GAAATTTACTGTGTGTGAGTCAGAGAGCCTGTGTGTTGGCACACGATGACTGAAAAAC 1585  
 QY 526 atcaagtagcagagatgagtagctgtagtaagaagtaggagagcagcaaggaaagca 585  
 DB 1586 ATTGCTATGTCGCCGGAAGAAATGTCACATGATGATGAGAAAGCTGTTAAGGAAGCC 1645  
 QY 586 aatgcgtatgattatataatgagagcttccataataatataatgagtaggagaaagaa 645  
 DB 1646 AATGCTATGATTTTATCATGAACCTAATAATTTGACACCTCTGTTGGAAGAGA 1705  
 QY 646 ggaagctaaatgtaggagcagaaacagagatcgcaatgctcgtgcttgaatcga 705  
 DB 1706 GGGGCCCCAGTGAAGTGGGACAGAAACAGAAATGCCATTGCTGGGCGCTGTTGCG 1765  
 QY 706 aaccocaaagattcgtatttgaatgagagctcgtccctggagtagcaagaagcaatcga 765  
 DB 1766 AACCCCAAGATTCTTCTGCTGATGAGCAACGCTACCTTGAGACAGAAAGTGAAGCA 1825  
 QY 766 gctgttcaagctcagctgagaaag 790  
 DB 1826 GTGGTTCAAGTGGCCCTGATTAAG 1850

RESULT 6

AX105057  
 LOCUS AX105057 4279 bp DNA PAT 30-APR-2001  
 DEFINITION Sequence 1 from Patent WO0123540.  
 ACCSSION AX105057  
 VERSION AX105057.1 GI:13921209  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 dog.  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE  
 1 (bases 1 to 4279)  
 Stocker,P.J., Stelmel-Crespi,D.T., Crespi,C.L., Reif,T.C. and  
 Paten,C.J.  
 TITLE  
 P-glycoproteins and uses thereof  
 JOURNAL  
 Patent: WO 0123540-A 1 05-Apr-2001;  
 GENTEST CORPORATION (US)  
 FEATURES  
 source  
 1..4279  
 Location/Qualifiers  
 /organism="Canis familiaris"  
 /db\_xref="taxon:9615"  
 17..3862  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC37718.1"  
 /db\_xref="GI:13921210"  
 /translation="MDPEGGRKGSAAEKFNKMKKSKNEKEKKPTVSTFAMFRYSN  
 WLDRLVLMVGMAIIGHGALPLMLVGFNMWDSFANAGISRNKTFPVILNISTNT  
 EIGHEDVHDYDELRLTLDVSKINEGIGKGFHRSIATFTFTGTVRGRKMLTL  
 VILASIPVILISAAIIMAKILISFPEKELIYAAAGAAVEVLAARIVIAFGQKEL  
 ERYKNLNEAKGILIGIKAITANISIGAAFLIIVASVALAWYGSITLVLSSEYSIGVL  
 TVEFSVLIGAFSIOGASIEAFAMARGAAVEIFKIIDNKPISIDYSKSGKPDNIG  
 NLEEFNHFESYPSRKEVILKGLKYOSSGOTVALVNGSGSGSTYVQLMQRLYDPTD  
 GMCICODIDRTINVRHLREITGVSOEPUVPAITIANINIGRENTYMDIEKAYE  
 ANAYDFIMKLPKFDITLVEERGAOLSGOKORAIARLVNPKIILIDATSAIDPE  
 SEAYVOALDARGRITIVIAHRLSTYRANADYAGEDDGVIVKGNDELMKEKGY  
 FKLTVMOTRNEIELEENATGESKESPDLEMSDGSGLIKRSTRSIHAOQODR  
 KLGTGKEDENVPVSEFWRLIKLNSTEMPYVVGICFALINGLOPASFISIFRIGI  
 FTREDEDEPTKRONSMSEVLEVLVIGISITFIFLOGFPFGKAGITRLRYMVRSM  
 LRODVSFEDDPKNTGATITTLANDAAOVKALISRLAVITONTANCTGIIISLTG  
 WQTLILLAVPIAIGVNEKMLSGALDKDELBGALIAEALENFTI  
 EOKREIYMAQSLQVYRNSLRKAHFGVSESITQAMMFYSYAGCFREAVYV  
 FQDLVLSAIVGMAVGVSSFAADYAKKVAANHIMIESKSPILDSYSPHGLK  
 NTEGNVTFENEVNYPTRPDIPIVLOGILEVKKGOTLIVGSGSGSTYVQLMQRLYDPTD  
 YDPLAGSVLIDGKEIKHLNVOMLAHGIYVSEDFILPECCSAENIYVGNR  
 IMAAKRANIHHFETLPEKNTYVGDGTLOSGOKORAIARLVNPKIILIDPE  
 TSAIDTSEKVVQALDARREGRTIVIAHRLSTIONADLIVFQNKVKEHHTHOL  
 LAQGIYFSMISVOAGAKR"

BASE COUNT 1294 a 834 c 1008 g 1143 t  
 ORIGIN  
 Query Match 27.6%; Score 324.2; DB 6; Length 4279;  
 Best Local Similarity 69.9%; Pred. No. 3,7e-64;  
 Matches 437; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 166 attattgataaagaccagafatagataactttccacagctgagatataactcgaatcc 225  
 DB 1124 ATAAATGCAATTAACCAAGCATTTGACAGCTATTGGAAGATGAGCATTAACCAAGCATTAAT 1183  
 QY 226 atagaaggaactgtagaatttaaaatgtcttcctcaatataccataagacatc 285  
 DB 1184 ATTAAGGGAATTTGGATTCGAAAATGTCTCACTCACTTACCCTTCGAAAAGAAATTT 1243  
 QY 286 aagattctgaaagtcyaaatcccaagaaatgaagtcgagagacagtcgccttgctggt 345  
 DB 1244 AAGATCTTAAGGGGTCTCAACCTGAAAGCTTCAGATGGGAGACAGTGGCCTGTTGGG 1303  
 QY 346 ctcaatgcaatggaagaagtagtagctctcagctctcagaggtatataatccgagat 405  
 DB 1304 AACACTGGCTCGGGAAGACAGACGCTGACAGCTGATGACAGGCTGTATGACCCACACA 1363  
 QY 406 gatgcttatacatgtagatgagaaatgacacagcagcttaaatgctgcgcatatcga 465









[illegible][illegible]

```
Db 1348 AACAGTGGCTGTGGAGAGACAAACAGTCCAGCTCATGACAGGCTCATATGACCCACCA 1407
Qy 406 gatgcttatacatggtgagatgaaatgacatcagaagcttaaatgctgcaattatcga 465
Db 1408 GAGGGGATGTCAGTGTGATGAGACAGGATATTAGACCATTAATGTAAGGTTTTCACGG 1467
Qy 466 gaccatattgagtggtgtaagaagcctgtttgttgcggagccaccctcagtaacat 525
Db 1468 GAAATCATTTGGTGTGTGTCAGGAACCTGTATTGTTTCCACCCGATAGCTGAAAC 1527
Qy 526 atcaaatgagcagagatattgtactgataagaagatgagagacgaagcaaggaagca 585
Db 1528 ATTGCGTATGAGCGCGTAANAATGTCACCATGATGATGATGAGAAACCTGTCAAGAGACC 1587
Qy 586 aatgcgtatgattatcatcgtgagttcctaataatataacattgtaggggaaaaa 645
Db 1588 AATGCTATGACTTATCATGTGAACCTGCTCAATTAATTGACACCCCTGGTGGAGAGAGA 1647
Qy 646 ggaagctcaatgagtgaggaagcgaagaacagagatgcgaattgctgtccttaagtca 705
Db 1648 GGGGCCCACTTGAAGTGTGGGCAAGAGACAGAGATGCCATTGACACCTGCTGTTGCG 1707
Qy 706 aacccaagaattctgatttgaatgagctacgtctgccttgattcagaagaagca 765
Db 1708 AACCCCAATCTCTCTGCTGATGATGAGCCACGCTGACCTTGACACAGAAACGGAAGCA 1767
Qy 766 gctgttcaagctgcacatgagaag 790
Db 1768 GTGGTTACAGTGGCTCTGTGATTAAG 1792
```

## RESULT 13

```
AF016535 4192 bp mRNA PRI 03-SEP-1997
LOCUS Homo sapiens P-glycoprotein (mdrl) mRNA, complete cds.
DEFINITION AF016535
ACCESSION AF016535
VERSION AF016535.1 GI:2353263
KEYWORDS
```

## SOURCE

## ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 4192)

Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I.,

Gottesman,M.M. and Roninson,I.B.

Internal duplication and homology with bacterial transport proteins

in the mdrl (P-glycoprotein) gene from multidrug-resistant human

cells

Cell 47 (3), 381-389 (1986)

87028230

2 (bases 1 to 4192)

Chen,G., Duran,G.E., Steger,K.A., Lacayo,N.J., Jaffrezou,J.P.,

Dumontet,C. and Sikic,B.I.

Multidrug-resistant human sarcoma cells with a mutant

P-glycoprotein, altered phenotype, and resistance to cyclosporins

J. Biol. Chem. 272 (9), 5974-5982 (1997)

3 (bases 1 to 4192)

Chen,G., Lacayo,N.J., Steger,K.A. and Sikic,B.I.

Direct Submission

Submitted (28-JUL-1997) Medicine, Stanford University School of

Medicine, Stanford, CA 94306, USA

Location/Qualifiers

1. 4192

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="7"

/map="7q21.1"

/cell\_type="drug-resistant uterine sarcoma"

/clone="8.1"

1. 4192

gene

```
/gene="mdr1"
/note="multidrug-resistance gene"
120..3959
/gene="mdr1"
/function="transporter protein"
/feature="mdr1"
/feature="multidrug-resistance"
/codon_start=1
/product="P-glycoprotein"
/protein_id="AAB69423.1"
/db_xref="GI:2353264"
```

## variation

659

/gene="mdr1"

/citation=[1]

/replace="c"

673..674

/gene="mdr1"

/citation=[1]

/replace="ct"

1351

/gene="mdr1"

/citation=[1]

/replace="c"

1428&lt;1429

/gene="mdr1"

/citation=[1]

/replace="ctc"

3551

/gene="mdr1"

/citation=[1]

/replace="c"

BASE COUNT 1243 a 797 c 1028 g 1124 t

ORIGIN

Query Match 27.2% Score 319.4; DB 9; Length 4192;

Best Local Similarity 69.4% Pred. No. 4.6e-63;

Matches 434; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

Qy 166 attattgataagaacacagatagataactttccacagctgagatataaactgatacc 225

Db 1218 ATTAATGATTAATTAAGCAGATGATGACATATTGACAGAGTGCGCACAACCATTAAT 1277

Qy 226 atagaagagacgtggaatttaaatgttcttcaattatccatcaagacatc 285

Db 1278 ATTAAGGAAATTTGCAATTTGCACTTCACTTACCATCTGAAAGAAAGT 1337

Qy 286 aagattctgaagctgagatcagaaatgaagctgagagagagagctgctgtcgt 345

Db 1338 AAGATCTTGAAGGCTGTAAGCTGAAGGTGACAGAGTGCGCAGACGCTGCTTGA 1397

OY	346	ctcaatgacgatgggaaagaatcacaggtacagctcttcacaaagttatgatccgat	405
Db	1398	AACAGTGGCTGTGGGAAAGCACAACAGTCCAGCTGATGACAGAGCTCTATACCCACA	1457
OY	406	gatggcttatacatggtgatgtaagaatcacacaagaagctttaaattgcygcatlatacga	465
Db	1458	GAGGGGATGGTCAAGTGTGATGATGACAGAGATATTAGACACATTAATGTAAAGTTCTACGG	1517
OY	466	gaccatattggaagtggttagtcaagaagctctgttttgctcggagccaccatcaagtaaat	525
Db	1518	GAATCATTTGGTGTTGGTAGTCAAGAACCTGTATTGTTTGGCCACACAGATAGCTGAAAC	1577
OY	526	atcaagaatgagcagaagatgatgacatgacatgaagaagatgaaagagcaagaaggaaga	585
Db	1578	ATTGCTATTGGCGCGGAAATGTCAACCATGATGATGATTGAAGAAAGCTGTCAAGGAAGCC	1637
OY	586	aatgcgtatgatattatcatcatgtaggagtttccctaaataattatcatatbgtlaagggaanaa	645
Db	1638	AATGCCATTATGACTTTATTCATGAACCTGCTCTCTTAATATTGGACACCCCTGTTTGAGAGAGA	1697
OY	646	ggaagctcaatagatggaggggcagaaacagaagaatcgcgaatcgtgcgtgccttaattcga	705
Db	1698	GGGGCCCAATTAGTGTGGTGGGCAGAAAGCAGAGGATCGCATTTGCAAGTCCCTGTTGCC	1757
OY	706	aaacccaaagatctgatatttagatgagctacagctctgcctcgtgattcagaagaagcaatga	765
Db	1758	AACCCCAAGATCCTCTCTGTGATGAGGCCACAGCTGAGCCTTGACACAGAAAGCGAAGA	1817
OY	766	gctgttcaagctgcacactgagaaga	790
Db	1818	GTGTTCAGGTGGCTCTGGATAGG	1842

RESULT	14
LOCUS	E02326 4378 bp RNA
DEFINITION	Multi drug resistance relating gene derived from human normal cells
ACCESSION	E02326
VERSION	E02326.1 GI:2170561
KEYWORDS	JP 1990100680-A/1.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Ueda,K. and Komano,T.
TITLE	1 (bases 1 to 4378)
JOURNAL	HUMAN NORMAL CELL-DERIVED MDR RELATED GENE Patent: JP 1990100680-A 1 12-APR-1990;
COMMENT	SUNTORY LTD OS Homo sapiens PN JP 1990100680-A/1 PD 12-APR-1990 PE 05-OCT-1988 JP 1988251475 PI UEDA KAZUMITSU, KOMANO TORU PC C12N15/12,C12N1/21,C12Q1/68; CC strandedness: Single; CC topology: linear; CC *source: tissue=Adrenal gland; CC *source: clone=SAM1132; FH *Key Location/Qualifiers FT 1..137 FT CDS 138..3980 FT mat_peptide 138..3977 FT /gene='Multi drug resistance relating gene' FT FT /gene='MDR1' FT FT 3'UTR Location/Qualifiers FT 1..4378 FT /organism='Homo sapiens' FT /db_xref='taxon:9606'
BASE COUNT	1315 a 818 c 1057 g 1188 t

## ORIGIN

Query Match	27.2%;	Score 319.4;	DB 6;	Length 4378;
Best Local Similarity	69.4%;	Pred. No. 4.7e-63;		
Matches 434;	Conservative 0;	Mismatches 191;	Indels 0;	Gaps 0;

Oy	166	atttctgataagaacccaglatagataacttctccacagctgggatataaacttcaatcc	225
Db	1239	ATAATTGATNTATNAGCCAGATATTGACAGTATTCAGAGTGGGCACAAACAGATAT	1298
Oy	226	atagaagaactgctggaattaaanaalyttcttcttcaattcaatccatcaagaacatcatc	285
Db	1299	ATPAGGGAATTTGGAAATTCAGAAATGTTCATCTCAGTACCCTTCGAAAAGAAAGT	1358
Oy	286	aagattctgaaagcttcgaatcttcagaatlaagctcggagaagaacagtcgcttgcgt	345
Db	1359	AAGATCTTGAAGGGCTCGAACTCGAAGGTGCACAGTGGGCACACAGCTGGCTGTGGA	1418
Oy	346	ctcatatgacagttggaagaagtaacgtagctcagctctctcaaggtatlatatccgag	405
Db	1419	AACAGTGGCTGTGGGAAGACACACACAGTCCAGCTATGACAGGCTGTATGACCCACA	1478
Oy	406	gatggtcttatcatcgttgatgtagaagatgacatcagaagctttaatgtgcgcatatcga	465
Db	1479	GAGGGATGGCTCACTGTGTGATGACAGAGATATTAGGACCTAAATGTAAAGTTTCTACGG	1538
Oy	466	gaccatattgaaaggttatacgaagccgttttcttgaggacaccatcaagtaacat	525
Db	1539	GAATTCATTGGTGTGGAGATCGAGAACCTGTTATTTTTCACACACGATAGCTTAAMAC	1598
Oy	526	atcaagaatggaacgagatgtgtgactgtatgaagaatgtagagagcagcaagaaagca	585
Db	1599	ATTGCGTATGGCCGTGAANAATGTACACATGAGATGTAGAAAGCTGTCAAGAACCC	1658
Oy	586	aatgcgtatgaatttatcatcgtgagtttccctaataaattaatatcatatgtagggaaaaa	645
Db	1659	AATCCCTATGACTTTATTCATGAAGAACTGCCATCAATTAATTTGACACCCCTGGTTGGAGAGA	1718
Oy	646	ggaagctcaaatagtagtagaggcagaaacagagatcgaaatgtgcgtgcgttagtctga	705
Db	1719	GGGGCCAGTTGAGTGGTGGGCCAGAAAGCAGAGAGATGGCATTGCCAGTTCCTGGTTGGT	1778
Oy	706	aaccccaagaattctgattttagatgtagagctacgcttcgccctggaatcagaagaacagta	765
Db	1779	AACCCCAAGATCTCCTCGTGGATGAGGCCACGCTGAGCTTGAGACACAGAAAGCAAGACA	1838
Oy	766	gctgttcaagctacgtcgggaag	790
Db	1839	GTGGTTCAGGTGCTCTGTGATTAAG	1863

RESULT	15
AR051647	
LOCUS	AR051647 4264 bp DNA
DEFINITION	Sequence 1 from patent US 5830697.
ACCESSION	AR051647
VERSION	AR051647.1 GI:5975011
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 4264)
TITLE	Slk1c,B.I. and Chen,G.
JOURNAL	P-glycoprotein mutant resistant to cyclosporin modulation
FEATURES	Patent: US 5830697-A 1 03-NOV-1998;
SOURCE	Location/Qualifiers
	1..4264
BASE COUNT	1263 a 807 c 1041 g 1153 t
ORIGIN	/organism="unknown"



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:39:46 ; Search time 293.81 Seconds  
(without alignments)  
2448.932 Million cell updates/sec

Title: US-09-873-409-12

Perfect score: 3177  
Sequence: 1 atggtcatctatcgtacacag.....atgcacagtcacgtgcagtga 3177

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_6/plodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/plodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/plodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/plodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/plodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/plodata/2/ina/Backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1124	35.4	4646	1 US-08-181-471-2	Sequence 2, Appli
2	1124	35.4	4669	6 5206352-3	Patent No. 5206352
3	1124	35.4	6505	2 US-08-793-610-5	Sequence 5, Appli
4	1124	35.4	9318	2 US-08-793-610-6	Sequence 6, Appli
5	1122.4	35.3	4264	2 US-08-784-649A-1	Sequence 1, Appli
6	1122.4	35.3	4264	2 US-08-784-649A-5	Sequence 5, Appli
7	1119.2	35.2	4669	2 US-08-583-276-18	Sequence 18, Appli
8	1114.4	35.1	4669	2 US-08-752-447-1	Sequence 1, Appli
9	1040.6	32.8	4233	3 US-09-120-513-1	Sequence 1, Appli
10	1040.6	32.8	4233	3 US-09-120-513-1	Sequence 1, Appli
11	856	26.9	2726	1 US-08-461-823-1	Sequence 1, Appli
12	426.6	13.4	4002	2 US-08-996-545-1	Sequence 1, Appli
13	426.6	13.4	4002	2 US-08-996-545-3	Sequence 3, Appli
14	426.6	13.4	4002	4 US-09-328-320-1	Sequence 1, Appli
15	426.6	13.4	4002	4 US-09-328-320-3	Sequence 3, Appli
16	386.6	12.2	4047	2 US-08-612-734B-1	Sequence 1, Appli
17	331.4	10.4	4224	1 US-08-612-521-1	Sequence 1, Appli
18	314.6	9.9	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	304.6	9.6	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	193.2	6.1	3927	2 US-08-996-644-1	Sequence 1, Appli
21	193.2	6.1	3927	3 US-09-352-552-1	Sequence 1, Appli
22	192.4	6.1	3924	2 US-08-996-644-3	Sequence 3, Appli
23	192.4	6.1	3924	3 US-09-352-552-3	Sequence 3, Appli
24	181.6	5.7	3909	1 US-08-232-537-1	Sequence 1, Appli
25	149	4.7	2061	4 US-09-061-764A-17	Sequence 17, Appli
26	145.2	4.6	2407	2 US-08-895-522-2	Sequence 2, Appli
27	145.2	4.6	2407	3 US-09-195-391-2	Sequence 2, Appli

28	145	4.6	5120	3 US-08-772-270A-6	Sequence 6, Appli
29	145	4.6	8370	2 US-08-488-706-1	Sequence 1, Appli
30	144.8	4.6	7721	3 US-08-772-270A-14	Sequence 14, Appli
31	141.2	4.4	1959	4 US-09-061-764A-4	Sequence 4, Appli
32	133.4	4.2	1201	3 US-08-961-083-189	Sequence 189, App
33	133	4.2	2244	4 US-09-061-764A-18	Sequence 2, Appli
34	130.4	4.1	4403765	4 US-09-103-840A-2	Sequence 1, Appli
35	128.6	4.0	2376	1 US-08-394-880B-1	Sequence 2, Appli
36	128	4.0	4403765	4 US-08-103-840A-2	Sequence 1, Appli
37	128	4.0	4411529	4 US-09-103-840A-1	Sequence 1, Appli
38	117.6	3.7	6143	1 US-08-612-521-3	Sequence 3, Appli
39	110	3.5	28804	2 US-08-592-874-1	Sequence 1, Appli
40	110	3.5	28804	3 US-09-096-942-2	Sequence 2, Appli
41	110	3.5	28804	3 US-09-096-867-2	Sequence 2, Appli
42	104.2	3.3	5889	1 US-08-463-092B-5	Sequence 5, Appli
43	104.2	3.3	5889	2 US-08-462-109A-5	Sequence 5, Appli
44	104.2	3.3	5889	2 US-08-460-907B-5	Sequence 5, Appli
45	104.2	3.3	5889	3 US-08-463-179A-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-08-181-471-2  
Sequence 2, Application US/08181471  
Patent No. 5641508  
GENERAL INFORMATION:  
APPLICANT: Li, Lingna  
APPLICANT: Lishko, Valery K.  
TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Thomas Fitting  
STREET: 12526 High Bluff Drive, Suite 300  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92130  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/181,471  
FILING DATE: 13-JAN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: ANT0029P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-792-3680  
TELEFAX: 619-792-8477  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4267  
US-08-181-471-2







Db	3159	aggaaacagaagtttgaaacatglatgctccagaggttcgaagttaccatagaaactctt	3212
QY	2138	ccaagaagaacagatattatggaagctgttatagcatccagcagtccttatatttg	2197
Db	3219	tgaagaaagacacacacttttggaattacaatttcctccaccgagaaatgatatttt	3278
QY	2198	cctatgcaagaggttcgatttggaagcctatttaattcaagctggaacgaatgaccgg	2257
Db	3279	cctatgctgagatgttcgcggttgtagagcctactgtgtgcacataaactatgaagcttg	3338
QY	2258	aggagctgttcaagttttactgcgaattgcacatgacatgagctatggccatcggaaacgc	2317
Db	3339	aggatgtctgttagatatttccagccgtgtccttggtgcacatgagccgttggaagta	3398
QY	2318	tcggtttgctctctgataatcccaagaccaaaccggggctgtgcacatgtttgcttgt	2377
Db	3399	gttcatttgcctcttgactagcccaaaagccaaatalcagaagcccatcatcatgta	3458
QY	2378	tggaaagaagaacccaatatagacagccgcagttcaagaaggaaagccgcagacatgtg	2437
Db	3459	ttagaaaaaaccccttgatgtacagctacagccagcgggaagccttaatgcccgaacatltg	3518
QY	2438	aaggaatttagaggtttcggaaagtcctcttcttctatccatgttcgcgcagatgtttca	2497
Db	3519	aaggaatatgtcacacttttggatggaaagtttatalcaacttcccaaccgacccgactccag	3578
QY	2498	tctctccgtgtctataccctcagattatgacgaggaagaagacagtaacattgtgtggagca	2557
Db	3579	tgtttcaggagactcgagccttgaggttgaaagaagccagacgctgctctgtgtggacgca	3638
QY	2558	gcggcctgttggaagaaagcacttcgttcaactctgcgcagagcctttagaccgccgtgcag	2617
Db	3639	gtgcctcttggaagaaagcacagttgttccagcctccggagcgggttctaagaccctltgcag	3698
QY	2618	gacaaagcgcgtttgtatcgtgtgtgatacgaagaaatgaaatgaaatgtaacagttgctccgttcc	2677
Db	3699	ggaaagcgcgtgtgtgatacgtgcgaagaagaataaagcgaacgaatgttcagttgcgtccgagac	3758
QY	2678	aaatagcaaatcgttctctcaagaagcctgtgtcttcaactgcagcatgtgcgaagaactcg	2737
Db	3759	acctggtcatcgtgtcccaagagcccatctcgtttgactgtgcagatattctgaagaacattg	3818
QY	2738	cctatgtgtgacaagaagccgtgtgtgtgcattagatgaaatgaaagaagccgcaaatgcag	2797
Db	3819	cctatggaagacaacagccgggtgtgtgtcacaagagaagatcgtgaggtgcagcaaaaggag	3878
QY	2798	caaatatccattcttttatttgaaagcttccctcctggaagaatacacaacaaagttggaactga	2857
Db	3879	ccaacataacagccttcaatcgatgcactgcgtctaataatatacgcctaagttagagagca	3938
QY	2858	aagagacacagcttctctgacgacgaagaacaagaactagctatttgcaaggctcttccc	2917
Db	3939	aaggaactcagctctctcgtgtgtgcgaagaacaagcattgcatagctcgttcccttgtta	3998
QY	2918	aaaaaaccacaatttatattgtttgtagtgccacttgagcccttgataatgacattgga	2977
Db	3999	gacagcccatatttctgttcttgtagaagccacgcctcagcctctgatacagaagaatgaa	4057
QY	2978	agggtgttcagcaatgcctctgataaagaacagagcaggaaggacatgctagtgatgcactc	3037
Db	4059	aggttgttccaagaagcctctggaacaagccagagaagcgcgacccgtgcatltgattgttc	4118
QY	3038	acaggtctctgtcaatttcagaaagcagatttgatgtgtgtctgcacaatgtaagaatga	3097
Db	4119	accgctgttccaccatccagatgacagacttaatatgtgtgtttagaattgccaagtga	4177
QY	3098	aggaaacaggaactatcaagaagcctcctgagaaatccggagacataatatttaagttagtga	3157
Db	4179	aggagcatgagcagatcatcagcgtgcgtgcacagaagaagcatctattttcaagtgtca	4237
QY	3158	atgcacagttcag 3169	
Db	4239	gtgtccaaagctg 4250	



Oy 1481 tctcgggagacattggtctctgttctaataatggaactggttcatccagttatttccatcatct 1540  
 Db 3951 TTGTTGTTGTTATTTTGGCATTATTAATGAGGCTCGACACGAGCATTTGCAATATATAT 4010  
 Oy 1541 ttcgcaaaattataaccatggtttggaataatgataaaccacattaaag---catgaty 1597  
 Db 4011 TTTCAAAAGATATAGGGGTTTTTACAAAGATTGATGCTTGAAACAAACACACAAATTA 4070  
 Oy 1598 cagaattattccatgatattcgttcatttgggttatttgggttatttgggttatttgggttattt 1657  
 Db 4071 GTAACCTGTTTACATTTGTTTCTGAGCCCTTGGAATTAATTTCTTTATTTATCTTTTCC 4130  
 Oy 1658 tgcagggattatttaacgcagagcgaggggaaatttaacgatgatataagacacttgg 1717  
 Db 4131 TTCAGAGTTTTCACATTTGGCAAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGG 4190  
 Oy 1718 ccttcaaaagccatgttatatacagatattgacctggtttgattgaagaagaaacagacag 1777  
 Db 4191 TTTTCGATCCATGCTCAGACAGAGATGTGATTGTTGATGACCTTAAACACCACTG 4250  
 Oy 1778 gaagcttgaacaacattatagccatagatataagacaattcaaggagacagatcca 1837  
 Db 4251 GACACTTGAACACTACAGGCTGCGCAATGATGCTCTCAAGTTAAAGGGCTATAGTTCCA 4310  
 Oy 1838 ggaatggcgtcttaacacaaatgcaactaacaatgggacttcaatatttcttcta 1897  
 Db 4311 GGCTTCTGTAATATCCCAAAATATAGCAAAATCTTGAGACAGAAATATATATCTTCA 4370  
 Oy 1898 tataatgatggagatgagacttctcgaatctgagttatgcttccagatcttgcgtgagag 1957  
 Db 4371 TCTATGTTGGCACTACACTGTTACTTACCAATTTGACATTTACCATTTGCAATACAG 4430  
 Oy 1958 gaatgatgaacocgacagcaatgactgatttgcacaaagaataagcaagaactaagc 2017  
 Db 4431 GAGTGTGTAATGAATAATGTTGCTGACAGACATGAAAGATTAAGAAACTAGAAAG 4490  
 Oy 2018 atgctggaagaatagcaactggaagcttggagaataatacgttactatagtgtcattaa 2077  
 Db 4491 GTGCTGGGAAGATCGCTACTACAAACATAGAAACTTCCGAACCGTGTGTTTGTGCTC 4550  
 Oy 2078 ggggaaagaccccttcgagcaaatgataatgaagatgcttcagacccaacagaaact 2137  
 Db 4551 AGGAGCGAAGATTGAACATATGATGCTCAGAGTTTGCAAGTACATACGAAACTCTT 4610  
 Oy 2138 cgaagaagacagagattatggaagctgttattgcaatcgaatcgaatgcttataattt 2197  
 Db 4611 TGAAGAAAGACACACTTTTGGAAATACATTTCTTCCACCCAGCAATGATGATTTT 4670  
 Oy 2198 cctatgcaagcaggttccgatattggagccttataatcaagcctggagcaatgcccag 2257  
 Db 4671 CCTATGCTGATGTTTCCGGTTTGAGGCTTACTTGTTGGCACATTAACCTCATAGACTTTG 4730  
 Oy 2258 agggcagttatagatttactgcaattgcatatagagctatggccatcgcaaaacgc 2317  
 Db 4731 AGGATGTTCTTATGATATTTTACAGCTGTGCTTTGTTGGCATGCGCTGGGGCAAGTCA 4790  
 Oy 2318 tctgttggctcctcgaatatacccaagccaaatcgggggctgagcatctgttgcctt 2377  
 Db 4791 GTTCATTTGCTGAGATATGCCAAAGCCAAATATCAGACGCCACATCATATCATATGATCA 4850  
 Oy 2378 tggaaagaacaaataatagacgcgcagtcgaagaaggaaagaaagccagacaatgtg 2437  
 Db 4851 TTGAATAAAACCCCTTTGATTTACAGCTACAGCACGGAAGGCTTAATGCGGAACATTTGG 4910  
 Oy 2438 aagggaatttagatttggagaagctcttcttcatccatgctgcccagaatgtttca 2497  
 Db 4911 AAGGAAATGTCACATTTGTTGTAAGTTGTAATTCACATATCCACCCGACGACATCCAG 4970  
 Oy 2498 tccctcgtcgtctatccctcagttatgagcgagaaagacagttacatttggggagca 2557  
 Db 4971 TGCTTCAGGAGACTGAGCTGGAGGTAAGAGGCGCACAGCGCTGCTGTGGGGCAGCA 5030  
 Oy 2558 gcgcgtgtgggaaagacacttctgttcaactctgcagagacttatagaaccccgtagag 2617

Db 5031 GTGGCTGTGGGAAGACACAGTGTCCACCTCTGAGCGGTTCTACGACCCCTTGGCAG 5090  
 Oy 2618 gaaagtgctgtttagtgtytgatgcaaaagaattgtaatgtaagtggtccgttcc 2677  
 Db 5091 GGAAAGTGTGCTGTGATGCAAGAAATTAAGCGACTGATGTTCAGTGCTCGAGCAC 5150  
 Oy 2678 aaatagcaactgtctcctcaagagccgtgtcctctcaactgcgaacattgttgaagaatcg 2737  
 Db 5151 ACCTGGGCACTGCTGTCCAGAGCCCATCTGTTTACCTCAGCATGCTGCGCAACATTG 5210  
 Oy 2738 cctatgtgacaacagcgtgtgtgtgcatatagatagatcaaaagaccgcaaatgacg 2797  
 Db 5211 CCTATGAGACACACCGGGGTGTGTACAGGAAGATCGTAGGGCGCAAGAGAG 5270  
 Oy 2798 caaatccaacttatttgaaggtctcctctgaagaataacacacaaagttggactga 2857  
 Db 5271 CCAACATACATGCTTATCATGAGTCACTGCTTAAATATATACCTAAAGTGAAGACA 5330  
 Oy 2858 aaggagcacagcttctcgtgggcagaaacaaagactagctatctgaaggtctctc 2917  
 Db 5331 AAGGAACACTGCTCTGTGGCCAGAAACACGATGCGCATAGCTGCGCTGTTA 5390  
 Oy 2918 aaaaaccacaaatttattgtttagatgagccacttcaagccctcgatataatgacagtga 2977  
 Db 5391 GACAGCTCATATTTTCTTTGGATGAGCCACGCTCAGCTGTGATACAGAAAGTGA 5450  
 Oy 2978 aggtgtgtcagatgccttgaagaagccagagcggaagaagacatgctgtgtgctc 3037  
 Db 5451 AGGTTGTCCAAGAGCCCTGAGCAAGACAGAAAGGCGCCACCTGATGTGATTCCTC 5510  
 Oy 3038 acaggtctctgcacattcagaacgacatttgaatgagtgtctgcacaaatgaaagataa 3097  
 Db 5511 ACCGCTGTCCACCATTCAGAAATGCACTTAATAGTGTTCACAAAGGCGAGACTCA 5570  
 Oy 3098 aggaacaagaactcatcaagagctccttgagaatcgaagacataatttaagtatga 3157  
 Db 5571 AGGAGCATGGCAGCATCAGACGCTGCGCAGAAAGGCAATCATTTTCAATGCTCA 5630  
 Oy 3158 atgcacagtcag 3169  
 Db 5631 GTGTCCAGGCTG 5642

RESULT 4  
 US-08-793-610-6  
 ; Sequence 6, Application US/08793610  
 ; Patent No. 5858744  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BAUM, Christopher  
 ; APPLICANT: STOCKING-HARBERS, Carol  
 ; APPLICANT: OSTERTAG, Wolfram  
 ; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
 ; NUMBER OF SEQUENCES: 6  
 ; ADDRESS: Nikaido, Marmelstein, Murray & Oram LLP  
 ; STREET: 655 Fifteenth Street N.W. Suite 330  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-5701  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/793,610  
 ; FILING DATE: 07-MAR-1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: DE P 44 31 973.8  
 ; FILING DATE: 08-SEP-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 195 03 952.1  
 FILING DATE: 07-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP95/03175  
 FILING DATE: 10-AUG-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Berman, Richard J.  
 REGISTRATION NUMBER: 39,105  
 REFERENCE/DOCKET NUMBER: P1614-7007  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)638-5000  
 TELEFAX: (202)638-4810  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9318 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA  
 US-08-793-610-6

Query Match 35.4%; Score 1124; DB 2; Length 9318;  
 Best Local Similarity 62.1%; Pred. No. 0;  
 Matches 1869; Conservative 0; Mismatches 1095; Indels 48; Gaps 4;

OY 200 ttgtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 259  
 DB 2596 ttgtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 259  
 OY 260 ctccaagaatgctctgctgctgctgctgctgctgctgctgctgctgctgct 319  
 DB 2656 ctccaagaatgctctgctgctgctgctgctgctgctgctgctgctgctgct 319  
 OY 320 ttgtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 379  
 DB 2716 ttgtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 379  
 OY 380 ttgctgttctcttgaatgcaagaatgcaagaatttggcataaaaggactag 439  
 DB 2770 ttgctgttctcttgaatgcaagaatgcaagaatttggcataaaaggactag 439  
 OY 440 ttgtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 499  
 DB 2830 ttgtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 499  
 OY 500 aaccacagatagatacatttccacagctgagataaaccatgacatagagaagactg 559  
 DB 2890 aaccacagatagatacatttccacagctgagataaaccatgacatagagaagactg 559  
 OY 560 tggaaatttaaaatgcttcttcaatatacacaagaacatataagaatctgaaag 619  
 DB 2950 tggaaatttaaaatgcttcttcaatatacacaagaacatataagaatctgaaag 619  
 OY 620 gtctgaatctcagaatgctgagagagagagagagagagagagagagagagagag 679  
 DB 3010 gtctgaatctcagaatgctgagagagagagagagagagagagagagagagagag 679  
 OY 680 ggaagagatag 739  
 DB 3070 ggaagagatag 739  
 OY 740 tggtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 799  
 DB 3130 tggtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 799  
 OY 800 tggtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 859  
 DB 3190 tggtaagtatacagaatctcaagaatgcaagaatttggcataaaaggactag 859  
 OY 860 ggaagatag 919

DB 3250 gtcgaagaatgctcaccatgagatgagatgagatgagatgagatgagatgagat 3309  
 OY 920 ttatacag 979  
 DB 3310 ttatacag 979  
 OY 980 gtcgaagaatgctcaccatgagatgagatgagatgagatgagatgagatgagat 1039  
 DB 3370 gtcgaagaatgctcaccatgagatgagatgagatgagatgagatgagatgagat 1039  
 OY 1040 tgaatttagatgag 1099  
 DB 3430 tgaatttagatgag 1099  
 OY 1100 gactgag 1159  
 DB 3490 gactgag 1159  
 OY 1160 ttgcaagtgag 1219  
 DB 3550 ttgcaagtgag 1219  
 OY 1220 atgctgaactaatgcaagaatgcaagaatgcaagaatgcaagaatgcaagaatgca 1279  
 DB 3610 atgctgaactaatgcaagaatgcaagaatgcaagaatgcaagaatgcaagaatgca 1279  
 OY 1280 aaaaagctgag 1332  
 DB 3670 aaaaagctgag 1332  
 OY 1323 aagaagaac 1386  
 DB 3730 aagaagaac 1386  
 OY 1367 -----agcagac 1420  
 DB 3790 -----agcagac 1420  
 OY 1421 ttctgaag 1480  
 DB 3850 ttctgaag 1480  
 OY 1481 ttctgaag 1540  
 DB 3910 ttctgaag 1540  
 OY 1541 ttgcaagaatgatacagaatgatacagaatgatacagaatgatacagaatgatacaga 1597  
 DB 3970 ttgcaagaatgatacagaatgatacagaatgatacagaatgatacagaatgatacaga 1597  
 OY 1598 cagaagaattatcagatgagatgagatgagatgagatgagatgagatgagatgagat 1657  
 DB 4030 cagaagaattatcagatgagatgagatgagatgagatgagatgagatgagatgagat 1657  
 OY 1658 tgcag 1717  
 DB 4090 tgcag 1717  
 OY 1718 ccttcaagaag 1777  
 DB 4150 ccttcaagaag 1777  
 OY 1778 gaggctgag 1837  
 DB 4210 gaggctgag 1837  
 OY 1838 ggaattgag 1897  
 DB 4270 ggaattgag 1897  
 OY 1898 tatatgag 1957  
 DB 4330 tatatgag 1957



OY	1958	gaatgattgaaacccgcagcaatgtaactgtgatattggccaacaagaataagcaagaactaaagc	2017
Db	4390	GAGTGTGTGAATATGAAATATGTGTCGTGGCAAGCAAGCACTGAAAGATTAAGAAAGCACTGAAAG	4449
OY	2018	atgcctggaagaatagcaactgaagactcttggagaataatacgaactatagtgactataacaa	2077
Db	4450	GTCCTGGGAAGATCCGCTACTGAAAGCATATGAAACACTTCGGAACCGTGTGTTCTTTGACTC	4509
OY	2078	gggaaaaagcccttcgagcaaatgtaabgaagatgcttcaagactcaacaagaataact	2137
Db	4510	AGGACACAGAGTTTGAACATATGATATGCTACAGAGTTTGGACGTATCATACAGAAACTCTT	4569
OY	2138	cgaagaagaacaaagtattatggaaagctgttaatgacctgaagccactgactatatttg	2197
Db	4570	TAAGGAACACACACTTTGGAAATTAATTTCCTTACCCAGGCAATGATATATTTTT	4629
OY	2198	ccatcgacgaagggttctgcatttggagccatttaatlcaagcttggacgaatgacccag	2257
Db	4630	CTATAGCTGATGTTTCCGTTTGGAGCCTACTTGGTGGCACATTAACATCATAGGCTTGG	4689
OY	2258	aggcactgttcaagtatttactcgcaattgcataatggagctatggcactatgggaaaaagc	2317
Db	4690	AGGATGTGTTGATTAATTTTACGCTGTGTCTTGTGGTGCATATGGCCGTGGGCAACTCA	4749
OY	2318	tcgatttggcctctgataatctccaagaagccaaatcggcgggctgagcatctgttggcctgt	2377
Db	4750	GTTTCATTTCCTCTCTGACTATATGGCAAAAGCCAAATATACAGAGCCACATCATCATATCA	4809
OY	2378	tggaaagaacaacaatatagacagccgcagtlcaagaagggaaaaaagcaacaacatgtg	2437
Db	4810	TTGGAATAAACCCCTTTGATTTGACAGCTACAGCACAGGAAGGCGTAATGGCCGAACACTTGG	4869
OY	2438	aaggagaatttagaagtttcgaaagctctcttcttcaaccatgtgcgcgaatgtttca	2497
Db	4870	AAGGAATGTACATTGTTGGTGAAGTTGTATTCAACTATCCACCCGACCCGAGATCCAG	4929
OY	2498	tctccgttgcttactccctcaagtatgttagagcgaggaagaagtagcaattgttggggagca	2557
Db	4930	TGCTTCAGGAGCTGAGCCTGTGAAGTGAAGGAAGGCCAGACGCTGGCTGTGGGGCAGCA	4989
OY	2558	gcgcctgttgggaaagacactctcgttcaacttctgcagaaactttagaaccctgtgcaag	2617
Db	4990	GTGGCTGTGGGAAGACACAGTGTGTCCAGCTCTGTGGAGCGGTTCATACGCCCTCTGGCAG	5049
OY	2618	gacaagtgccttcttgatgtgtgtgatagcgaagaatgtaatgtaagtgagctccgttccc	2677
Db	5050	GGAAATGCTGCTTGTATGGCAAGAATATAAGGACATGATGTTCACTGGCTCGGACAC	5109
OY	2678	aaatagcaactcgttccctcaagagccgtgtccttcaactcgaactgtgagagaacatcg	2737
Db	5110	ACCTGGGCAATCGTGTCCCAAGAGCCCAATCCTGTTTGACTGCAGCATTTGCTGAGAACTTG	5169
OY	2738	cccatgtgtacaacacagccgtgtgtgtgcataatgataccaagaagacgcgcaaatcgag	2797
Db	5170	CCATGTGAGACAAACAGCCGGGTGTGTGCACAGGAAGAGATGTGAGGGCAGCAAAAGAGAG	5229
OY	2798	caaatatcaatcttattatgaaagctccctcgaagaataacaacaacaagtgtgactga	2857
Db	5230	CCAACATACATGCTTCATCAGTCACTGCTATTAATATATATAGCACTAATAAGTAGAGACA	5289
OY	2858	aaggagaacagcttcttgcgcgcgaagaanaaagacagactatgtgaagggctcttccc	2917
Db	5290	AAGGAACCTGACCTCTCGGTGGGCCAGGAACACGCATATGCCATATAGCTGTGCCCTTGTTA	5349
OY	2918	aaaaaccaaatttatcttggatbgaagccacttaagccctcgtaatatgacagtga	2977
Db	5350	GACAGCCTCATATTTTGGCTTTTGGATGAAGCACAGCTCAGCTCGGATATACAGAAAGTGAA	5409
OY	2978	aggltgttcagcatgcctctgataaagccagacgcygaagagacatgctctagtgatcactc	3037
Db	5410	AGGTTTGTCCAAGAGCCCTGTGACAAACCCAGAGAGGCCGACCTCGCATTTGTGATTGGCTC	5469

QY	3038	acaagcctctcgcaatlcgaacgcgagatttgtagtggllcttcgcaaatggaaataa	3097
Db	5470	ACGCCCTGTCACCATCCAAATGCAGACTTATACTGTGTTTCAGATTGGCAGATCA	5529
QY	3098	aggagaacgaactcatcaaagctlctctyagaatlcgcgaacatatatltaagttagga	3157
Db	5530	AGGAGCATGGCACCATCACCACTCTGTGGACAGAAAGCATCTATTTTCCAATGTCA	5589
QY	3158	atgcacagctcag	3169
Db	5590	GTCGCCAGGCTG	5601

```

RESULT      5
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: 'Sikic, Branislav I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-784-649A-1

Query Match      35.3%; Score 1122.4; DB 2; Length 4264;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 1868; Conservative 0; Mismatches 1096; Indels 48; Gaps

QY      200      tttaagatatacacagaatctcaagaatgacaaagatttggcatalaaagaagactatag 259
           ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      959      ttgaagatgacaaacaaattttgaagaagatgaagaaatttggaattgaagaaactatga 1018

QY      260      cttaacaaatgtctcttgctgtgctgtgactcttcttataagatgaaacttggaacttggt 319
           ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      1019     cagcccaattttttatagaagtgctgctttctcgtgattgacatcttattgcttg6cct 1078

QY      320      ttgtgataggaacctcttgatctcttaagtggaagaacctgatatacacaggaactgttc 379
           ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1079     tctgtataggagaccacttggtgctctctcagggaa-----attattctattggacacagttac 1132

```

Db	1133	TCACGTATTCTTTCTGTATTATTTATTTGGGGCTTTTATGCTTGGAGAGCAGCATCTCCAAACA	11932
Qy	440	ctgaaaccttcgcaatagccccgagagctgccttcataatttccaggcttatcataaga	499
Db	1193	TTGAGACATTTGGCAATGCAAGAGGAGAGCTTTATGAAATCTTCAAGTAAATTATGATAATA	1252
Qy	500	aaccagatagataaactttccacagctgatatataaacctgatatccataaagyaactg	559
Db	1253	AGCCAAATTTGACAGCTATTTCGAAAGTGGGGCACAAACACAGATAAATTTAAGGGAATTT	1312
Qy	560	tggagatttaaaatgcttcttcaatataatccataagaccatctataagattctgaag	619
Db	1313	TGGAAATTCAGAAATGTTCACTTCAGTTACCTCCATCTCGAAAAGAAAGTTAAAGTCTTGAAGG	1372
Qy	620	gtctgaatccagaattaaagtctctgagagagacagctgcctttgtctgtccatgagcag	679
Db	1373	GCCGAAACCTGAAAGGTGCAGAGTGGGGACACGGTGGCCCTGGTTGAAACAGTGGCTGTG	1432
Qy	680	ggaagagtaacgtagtccagctctctgcagaaggctaatatgatccagatgatgtgttata	739
Db	1433	GGAGAGCACACACAGTCCAGCTGATGCAAGAGCTGTATGACCCACAGAGGGAGATGCTCA	1492
Qy	740	tgtgtatgtaagatgacatcagagctttaaattgtcgcgcatatacagaaacataltgag	799
Db	1493	GTTGTTATGACAGAGATATTAGGACCATTAAGTTAAAGTTTTCAGGGGAAATCATTTGGTG	1552
Qy	800	tgtgtatgaagaagccgtgtttcttgagagaccatccagtaacatatcaagtatgac	859
Db	1553	TGGTGAATCGAGAACTGTTATTTTGGCCACACAGATAGCTAAAAACATTGGCTATGGCC	1612
Qy	860	gagatgtagtgcattgalyaagagatgagagagcagcaagaggaaacaaatgcgtatgat	919
Db	1613	GTCGAAATGTCACCATGATGATGAGATTTGAGAAAGCTGTCAAGAGAACCCATGCTATGACT	1672
Qy	920	ttatcattgaggttctctataaaatttaatacatattgttagggaaaaagagctcaatga	979
Db	1673	TTATCATGAAGAACTGCCCATTAATTTTGAACACCTGTGTTGGAAGAGAGGGGCCAGTTGA	1732
Qy	980	gtgagaggcagaacaagaagatctgcaatttgcgtgccttagttcgaaaccccaagatgc	1039
Db	1733	GTTGGTGGGCGAAGACAGAGGATGCCATTGTCAGTGTCCTGGTTGCAACCCCAAGATCC	1792
Qy	1040	tgaattttagatgaggtctacgtctgcctctgattccagaagaagtcagctgttcaactg	1099
Db	1793	TCTCGATCGAATGAGGCCACCTCAGCTTGGACACGAAAGCAACAGACAGTGTCTCAGTGG	1852
Qy	1100	cactgsgaagcgcgagcaaaagtcggaactacatcgttggtaagacacgacttctacta	1159
Db	1853	CTCTGGAATTAAGGCCACGAAAGGTCGGAGCACCATTTGTGTACCTCATGCTTGTCTCAG	1912
Qy	1160	cttgaagtgcagatttgattgtgtacccctaaagatgaaatgctgtgcggagaaggaagac	1219
Db	1913	TTTCTGAATGCTGACGTATCGTGTGTTTGGATGATGCACTTGTGGGAAAGGAATTC	1972
Qy	1220	atgctgtgaactatgycgaacacgcgagctctatatattacactgtgtatgtcaagagata	1279
Db	1973	ATGATGAACATCATGAAGAAAGAAAGGCAATTACTTCAAACCTTGTCAATGCAAGACAG	2032
Qy	1280	aaaaagctgtagaacaagatgtagtcaatgacatattcta-----ctga	1322
Db	2033	GAAATGAAGTTGAATTTGAAATAATGACAGCTGATGAAATCCAAAGTGAATTTGATGCCTTGG	2092
Qy	1323	aagaagaaccaactctctctctgnaactctgtgaagagcatca-----	1366
Db	2093	AAATGCTTTCAAAATGATTCAAGATCCAGTCTTAATAAGAAAAAGATCAACTGTFAGAGTG	2152
Qy	1367	-----agtcagacttcatgcaagcgtgaggaatccacccaactctaagagataagtc	1420
Db	2153	TCCGTGGATCACAAGGCCAAGACAGAAAGCTTAGTATCCAAAGAGCTCTGGATGAAATA	2212
Qy	1421	ttctctgaagtcctctataaaaaatttaaagttaaacgaacgctgaatgagccttcttg	1480

Db	2213	TACCTCCAGTTTCCTTTTGGAGATTAATGACGTAAATTTAACTGAATGGCTTAATTTTG	2272
Qy	1481	tctccgggagacatttgcttcctcgtctctaaatggaacgtgtcacaglatcttccalcact	1540
Db	2273	TTGTGGGTATTTTGTGGCATATTATAATGAGAGCGCTCGAACACCATTTGGCAATAATAT	2332
Qy	1541	ttgcataaaattataccaatgtttgggaaataatgaataaacccatlaag---catgatg	1597
Db	2333	TTTCAAGATTAATATGGGGGTTTTTACAAAGAAATGATATATCTGAAACAAACACGAGATA	2392
Qy	1598	cagaaattatccaigatatctgcgaatttgggtgttaatttgccttgctgcagtatcttca	1657
Db	2393	GTAACCTGTTTTCCTCTATTTGTTTCAGCCCTTGGAATTAATTTCTTTTATTATCATTTTTC	2452
Qy	1658	tgcagggatcttcttaacgcagacaggggaaattttaaagatgaagatlaagacacttgg	1717
Db	2453	TTTCAGGTTTCCACTTTTGGCAAGAGCTGAGAGATCTCTCACCAAGGGGCTCGATACATNG	2512
Qy	1718	cccttaaaagccatgttatatccagatatgtcctgtgttgaatgaaaagaaacagacag	1777
Db	2513	TTTTCCGATCCATCTCAGACAGAGATGTGAGTTTGTTATGATACCTTAAAAACACACCTG	2572
Qy	1778	gaagcttgacaacaatatcttgccatgatataagcaaaattcaaggagcaacagtttcca	1837
Db	2573	GAGCATGTACTACACAGGCTCGCCAAATGATGCTCTCAAGTTAAAGGGCTATAGGTTTCA	2632
Qy	1838	ggaattgagcttcttaacaaatgaacatacaatggagacttccagtatccattcctctta	1897
Db	2633	GGCTTGGCTGTAATTACCCAGATATAGCAAAATTTGGACAGAGGAATTAATATCTTCCA	2692
Qy	1898	tatatgattgggagatgaatccttcgatacttcgattgtctccagtaacttgcgttgacag	1957
Db	2693	TCTATGTTGGCACTAACACTGTATCTCTTACCAATTTGACCCATCATTTGCAATAGCAG	2752
Qy	1958	gaatgatgaacccgcagcaatgacatgagattgttccaaagaagatgaacgaagactaagc	2017
Db	2753	GAGTTGTTGAATGATAAATGTTGTGTGGACAAACACTGAAAGATATGAAGAACTAGAG	2812
Qy	2018	atgcctgaaagatgatccaactggaagccttctgagagataatagtactaagtgtcataaaca	2077
Db	2813	GTGCTGGGAATATGGTACACTGATGAAGCAATATAAATCTCGAACCCTGTGTTTGTGACTC	2877
Qy	2078	gggaaaaagcccttcagcaaatgtatgaagagatgtcttcaagatcaacacagaaatacc	2137
Db	2873	AGGACACAGATTGTAACATATGATGCTCAGAGTTTGCAGGACCATATCAAAACTCTT	2932
Qy	2138	cgaagaagaacacagataatttggaaagctgttatgcatcatcagccatgcttataatatltg	2197
Db	2933	TGAGGAAAGCACACACTCTTTGGAAATTAACATTTTCTTCCACCCAGGCAAGATGATTTTTT	2992
Qy	2198	cctaagcagcagggcttcgattcttggagccatttaattcaagcttgagcagatgagaccag	2257
Db	2993	CCTATGCTGGATGTGTTCCGGTTTGGAGACCTTACTTTGGTGCCATTAACATCATGACTTTG	3052
Qy	2258	aaggacatgtcaatagtcttcttaacgtcaatltgcataatgagactatggtccatcggaaaaagc	2317
Db	3053	AGGAGTCTCTGTATGTATTTTTCAGCTGTGTCTTTTGGTGCAATGGCCGTGGGGGAATGCA	3112
Qy	2318	tcgttttgctctcgtgaatatccaagaacaaatccggggctgcgcatactgttttcctgtc	2377
Db	3113	GTTTCATTTTGTCTCCGACTATGCCCAAATATATCACAGGCCCATCATCATCATGATGATCA	3172
Qy	2378	tggaaaaagaaccaaataatagacagccggcagtcgaagaaggaaaaagccacacatatgt	2437
Db	3173	TTGAAAAAACCCCTTTGATTGTACAGCTACAGCAGGAAGGCTTAATGCCGAACACATTTGG	3232
Qy	2438	aagggaattttagatlttctcgaagaagctccttctcttcatccatgtcgcacagatgttttca	2497
Db	3233	AAGGAATATGTACATTTTGTGTGAAGTTTATTAACATATCCACCCGAGACCGGACATCCAG	3292
Qy	2498	tctctcgttgcttatccctcagtaattgagcaggaagaagacagatgacatttgtggggagca	2557
Db	3293	TGCTTTCAGGAGCATAGCCTTGAGGTGAAGAAGGCCACGACGCTGTGCTGTGTGGGACACA	3352



Oy	2558	gcggcgtctggagaaagacacttcgltcaactcttgcgaagacttatatcccgylgcaag	2617
Db	3353	GTGGCTTGTGGAAAGAGACAGTGTCGTCACACTCCTGGAGCGTTCTACGAACCCCTTGGCAG	3412
Oy	2618	gacaacgtcgttttgattgtygtlgtalgcaaaagaatltgaatlgtaaglygtccgttccc	2677
Db	3413	GGAAGAAGTGCTGTATTGATGGCAAGAANAATAAAGGAGACTGMAATGTTCAHGGCTCCGAGAC	3472
Oy	2678	aaataagcaatcgyttcctccaagaagaccgtgctcttccaactgcgaagctbtctgaagaacctg	2757
Db	3473	ACCTGGGCATCTGTGCCACAGACCCTCATCTCTGTTCACTCGACGACTTCTTAGAACATG	3552
Oy	2738	cctaaggtagcaaacagccgcygtgtgtgccatatgatgatgaatcaagaagccgaatgcag	2797
Db	3553	CCTATTGGAGACAACAGCGGGGTGTGTCTCACAGGAAGAGATTGTGAGGGCAGCAAMAAGAHG	3592
Oy	2798	caaatalcatcttcttatttgaaagtccctcgtgagaatatacacacaaagtltgactga	2857
Db	3593	CCAACATACATAGCCTTCATCGATGATGCACCTGCCTAATTAATATGACCTAAAGTAGAGACA	3652
Oy	2858	aagagagacacagcttctcgtgagcgccaaaaaaaagactagtctatlgtcaaggtctcttccc	2917
Db	3653	AAGAACTCAGCTCTCTGGTGGCCAAAGAACACAGCATTTGCCATTCCTGTCCTTGTTA	3712
Oy	2918	aaaacccccaaatttatttctltgtalgtagagccacttcaagccctcgatataatgacagtga	2977
Db	3713	GACAGCCTCATATTTCCTTCTTTGTGATGAAGCCACGTCACCTGTGATACAGAAAGTGAAA	3772
Oy	2978	aggtgtgttcagcatgcctcttgtataaagccagaagacgggaagacatgctagtgtgtcactc	3037
Db	3773	AGGTTGTCCAAAGAGCCCCTTGACAAAGCCAGAGAAAGCCGACCCGCACTTGATGTGATTC	3832
Oy	3038	acaagcctctcgtcaattcagaacagccagatttgatagttgtctgcacatgtgaagataa	3097
Db	3833	ACCGCCTGTGCACCATTCACAAATGCAGACTTAATAGTGTGTTTACAGATGGCAGAGCA	3892
Oy	3098	aggaaacagaactcalcaagaagctccctcgtgagaatcgaacatcataatttaagtlagtga	3157
Db	3893	AGGAGCATGGCAGCCATCACACAGCACTGCTGGCACAGAAAGCATCTATTTTCAATGATGCA	3952
Oy	3158	atgcacacagtcag 3169	
Db	3953	GTGTCCAGGCTG 3964	
 RESULT 6 US-08-784-649A-5 Sequence 5, Application US/08784649A Patent No. 5830697  GENERAL INFORMATION: APPLICANT: SIKIC, Brainlimr I APPLICANT: Chen, Gang TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO NUMBER OF SEQUENCES: 5 CYCLOSORIN MODULATION CORRESPONDENCE ADDRESS: ADDRESSEE: Fish & Richardson STREET: 2200 Sand Hill Road CITY: Menlo Park STATE: CA COUNTRY: USA ZIP: 94025  COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PatentIn Release #1.0, Version #1.25 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/784,649A FILING DATE: CLASSIFICATION: 435 ATTORNEY/AGENT INFORMATION:			

[illegible]

Db	1733	GTGGTGGGCGAAGGACGAGGATGCGCATTTGCAACGTCCTGGATTGGCAACCCCAAGATCC	1792
Qy	1040	Tgaattttagatggaggtctacgctctgcctcttgatctcgaaagcaagtcagctcttcaagctg	1099
Db	1793	TCCGTGCTGGATGAGGGCCAGCTGACGCTTGGACACAGAAAGCAACATGTTGACAGTGG	1852
Qy	1100	cacttgagaagcgagcaaaagtcggaatacaacgcgtgtagcacaccgacttctacta	1159
Db	1853	CTCTGGATTAGGGCCAAAAAGGTCGGACCAACCATTTGTGTAGCTCATGTGTTGTCACAG	1912
Qy	1160	ttcgaagtgcaggtttgattgtgacccctaagaatgynaatgctgycggaagaagagac	1219
Db	1913	TTTCGTAATGCTGACGTCATCCTGCTGTTTGGATGATGGAGTCATTTGTGAGAAAGCAATC	1972
Qy	1220	atgcgtgacaatcagcgcaaaacagagtcatactatactcaactgtgtagcacaagatatta	1279
Db	1973	ATGATGAACCTCATGTAAGGCAAAAAAGGATTCCTTACTTCCAACTTGTCAATGCAAGACAG	2032
Qy	1280	aaaaagctgatacagaatgagtgatcaatgacatattcta-----ctga	1322
Db	2033	GAATTCGAAGTTCGAATTAAGAAAAAGCGATGATGAATTCAAAAAGTGAATTCGATGCTTGG	2092
Qy	1323	aagaagaaccaactcactccctctgacctctgtgaagacatca-----	1366
Db	2093	AAATGCTCTCAATATGATTCAGATCCAGTCTAATTAAGAAAAAGATCAACTGTAAGAGTC	2152
Qy	1367	-----agtcagacttcatatgacaagcgtgagaatcccccacactaagaagataagtc	1420
Db	2153	TCCCTGGATTCACAAAGGCCAAAGACAGAAACCTTAGTACCAAGAGGCTCTGGATGAAAGTA	2212
Qy	1421	ttctctgaagctctctctattaaaaattlaaagtltaacaagcctgaatgaccttltg	1480
Db	2213	TACCTCCAGTTTCCTTTGGAGGATTAATGACCTAAATTTAATCAGATGGCCTTAATTTTG	2272
Qy	1481	ttctgggggaacttggtctctgtcttaaaaggaactgttcacaglatittccatcatct	1540
Db	2273	TTGTGGTGTGATTTTGGCCATTATTAAGGGAGGCTCGCAACACATTTCGCAATATAT	2332
Qy	1541	ttggaanaaatataaacaatgttctggaataatgataaacaacacatlaag---catgatg	1597
Db	2333	TTTCAAAAGATTTATRGGGTTTTCACAAAGATTTGATGATCTGTAAACAAACAGACAGATTA	2392
Qy	1598	cagaanaatttccatgatactcgtcaattcttggtgtgatactgtgtcttgatctca	1657
Db	2393	GTACACTTGTTTTCACATATGTTTCTAGCCCTTGGAAATTAATTTCTTTATTCATTTTTC	2452
Qy	1658	tgcaggaatacttlaecgcgcagagcgagggaaattlaacgaatgagatlaagaaacttgg	1717
Db	2453	TTTCAGGGTTTACATTTTGGCAAAAGCTGGAGAGATCTCTCACCAAGGGCTCCGATATACGG	2512
Qy	1718	ccctcaagcagatgtatatctcagagatatgctcgtgtttgataaagaagaaagaaacag	1777
Db	2513	TTTTTCGATCCATGCTCAGACACAGAGATGTAGTTGGTTTGATGATACCTTAAAAAACCCACTG	2572
Qy	1778	gagggcttgacaacaatatactgacatagatataagcaaaatlcaaaagagcaacagttlca	1837
Db	2573	GAGCATTTGACATACAGGCTGGCCAAAGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCA	2632
Qy	1838	ggattgagcgtcttaacacaanaatgacaactaacaatgaggaacttcagttatcatcttccta	1897
Db	2633	GGCTTGGCTGTAATATACCCACAAATATATGCAAACTTGGGACAGGAATATATATCTTTCA	2692
Qy	1898	tatagtagggagagagactctcigtatctcigatagatatactcagtaacttcgltgcgtacag	1957
Db	2693	TCCTATGGTTGGCACTACACTCTTACTCTTAGCAATTTGTATCCCAATTCATGCATATAGAG	2752
Qy	1958	gaatgaatgaaacccgcagcatgactacgtgatttggccaacaagaataagcaagaactiaagc	2017
Db	2753	GAGTGTGTTGAATGAATAAATTTCTCTGGACAAACACCTGAAGAAATATTAAGAAAGACTAGAG	2812
Qy	2018	atgcttggaagaagatagcaacttgaaagcttggagaataatacgtatactatagtgatcaatcaa	2077

Dd	2813	GTGCTGGGAGATGCGCTACTGAGCAATAGAAAACTCCGAACCGTGTCTTGACTC	2872
Qy	2078	gggaaaaagcccttcgagccaatglatgaaagatgcttcacgtcacaacagaataact	2137
Dd	2873	AGGAGCAGAGATTGGAACATATGTATGTCACAGGTTTGCAGTACCATTAAGAAACTCTT	2932
Qy	2138	cgaaagaagcacagatlatatggaagctgttatgcatctcagctccattatatttgy	2197
Dd	2933	TGAGGAAAGCACACATCTTTGGATTACATTTTCCCTTCACCCAGGCAATGATGTAATTTT	2992
Qy	2198	ccctatgaagcagggttcgatttggatggccattatlaaactcagctbgagaaagcccg	2257
Dd	2993	CTTATGCTGGATGTTTCCGTTTGGAGCCCTACTTGCTGGCATATAACATGACCTTTG	3052
Qy	2318	tcgcttcgcctccgtaatlctccaaagccaalcggyggctgcgcatcgtlttcctgt	2377
Dd	3113	GTTCATTTTGCTCCCTGACTATGCCAAAGCCAAATATACACACCACCAATCATATGATACA	3172
Qy	2378	tggaaagaagaacccaataatgacagccgagtcagaagaaggaagaagcagacacatgy	2437
Dd	3173	TTGAAAAAACCCTTTGATTGACAGCTACAGCACGGAAGCCTTAATGCGAACATATGG	3232
Qy	2438	aaggaaatltagatlttcgaaagctctctctcatalccatcgaatgcgcagaatlttca	2497
Dd	3233	AAGGAATGTGCATTTTGGTGAAGTTGTATTCATCATCTCCACCGGACGGACATCCAG	3292
Qy	2498	tcctccgtgagcttaccctcaatgatgagcgagaaagacagtagcatltgtggggagca	2557
Dd	3293	TGCTTACAGGAGCTGAGCCTGGAGGTGAGTAAGAGGGCCAGACGCTGCTGTGGGGCACCA	3352
Qy	2558	gcgcgtctgaggaaagaagcatctcgttcaactcttgagaagacttgaaccccgycag	2617
Dd	3353	GTGGCTGGGGAAGACGACAGTGTCTCAGTCTCTGGAGGGTTCTACGACCCCTTGGCAG	3412
Qy	2618	gacaaagtgcgtttgattgagtggtgatlgaacaaagaatlgaatgtaagtgatccgtcc	2677
Dd	3413	GGAAGATGCTGCTTGATGAGCGCAAGAAATAAAGGCACTGAATTTTCAGTGGCTCCGAGAC	3472
Qy	2678	aaatagcaatcgttccctcaagaagcgtgtgctcttcaactgacagcatlgtcgaagaac	2737
Dd	3473	ACCTGGGCACTGTGTCCAGAGAGCCCATCTGTTTACTGACAGCAATWGCTGAGAACATYG	3532
Qy	2738	ccctatgttgacaacagccgtgtgtgtgcatttagatgatatcaagaagccgcaatcgag	2797
Dd	3533	CCCTATGGAGACAAACGCCGGGTGGTGTCAAGAACAAAGTTGTGAGGCGACGCAAAAGGAG	3592
Qy	2798	caaatatccatctcttatttgaagtgctccctcgtgagaatatcaacacaaagttgagca	2857
Dd	3593	CCAACTATCATGCTTATATGAGTCACTGCTTAATAATATATGCACTAAAGTAGACACA	3652
Qy	2858	aaggagcacagcttctcgcgcgcagaaacaaagaactagatctatgcgaagcgctctcc	2917
Dd	3653	AAGGAACTCACACTCTGTGHTGGCCAGAACAAACGATTCGCCATAGCTGTGCTGTGTA	3712
Qy	2918	aaaaaaccaaaatttatgtgtgatlgaagcgcaactcagacccttgataatgacagtga	2977
Dd	3713	GACAGCCCTCAATATTTTCTTTTGGATGAACCCACGTCACCTCTGGATACAGAACTGTAA	3772
Qy	2978	aggttgattcagcatcgccttgatlaaagccagagccggaaagagcatcgcctagtgatctc	3037
Dd	3773	AGGTTGTCCAAAGAACCCCTTGACAAAGCCAGAGAGGCGCACCTGCAATTGTGATTGCTC	3832
Qy	3038	acagagctctctgcaatctcagacgcagatltgatlgtgtctgcaacatggaaagataa	3097
Dd	3833	ACCGCTGTCTCCACCATCCAAATGACAGACTTAATAGTGTGTTTCAGAAATGCGACAGTCA	3892
Qy	3098	aggaaacaagaatcatcaagaagctccctgagaatctcgaacatatatttlaagtlagtga	3157
Dd	3893	AGGAGCATGAGCACCATTCACACTGCTGGCACAGAAAGCATCTAATTTTTCATATGCTCA	3952





```

: TELEFAX: 312-913-9808
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4669 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..424
: NAME/KEY: CDS
: LOCATION: 425..4264
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 4265..4669
: US-08-752-447-1

```

```

Query Match          35.1%; Score 1114.4; DB 2; Length 4669;
Best Local Similarity 61.9%; Pred. No. 1.5e-314;
Matches 1863; Conservative 0; Mismatches 1101; Indels 48; Gaps 4;

```

```

OY 200 ttgtaagatatacacagaatctcaagaatgcaagaatttggcataaaaaaggactatag 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1245 TTGAAAGATGACAAACAAATTTAGAAAGCTTAAAGAAATGGATAAAGAAAGCTATTA 1304
OY 260 ctcaaaagtgtctctgtggtgctgtactctcttaagtaagtaagtaagtaagcttgc 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1305 CAGCCATATATTTTATAGGGTGCTGCTTCTGCTGATCTATGCACTTATGCTGTGGCCT 1364
OY 320 ttgtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1365 TCGGATAGGACACACCTTGCTGCTCAGGGA-----ATATTCATTTGACAAAGTAC 1418
OY 380 ttgctgttctcttaagtgtatccatagcagttatgcatgttgagcagcagtcctact 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1419 TCACCTGATTTCTTTCGTATTAATTTGGGCTTTAGTGTGGACAGGCAATCTCAAGCA 1478
OY 440 tttaaaccttcgacataagccgagagctgccttcataatttccaaagttatgataaga 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1479 TTGAGATTTTGCAAATGCAAGAGGACAGCTTATGAATCTTCAAGATATATGATATA 1538
OY 500 aaccagatagataactttccacagctgataataaactgaatccatagagaagactg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1539 AGCCAAATATGACAGCTATCGAAGAGTGCGACAAACAGATATATTAAGGGAATTT 1598
OY 560 tgaatttaaaaatgttcttcaatatccatacaagacccatcaagaatctgaag 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1599 TGGAAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAGAAAGATTAAAGATCTTGAAGG 1658
OY 620 gtctgaatctcaagaatgaagcttgagagacatcgcttgctgctcaatgcaagtg 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1659 GCGTGAACCTTAAGGTGAGAGTGCGGACGAGCGGTGGCCCTGGTTGGAACACATGGCTGTG 1718
OY 680 ggaagtaagctgtaagcttcagcttcgcaaggttataatgatacggatgagttatca 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1719 GGAAGAGCAACAACAGCTCAGTATGCAAGAGGCTATGACCCACAGAGGAGGATGTCA 1778
OY 740 tggtagatgagaatgacatcagaagcttcaaatgtgcygcatatcgagacatattgag 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1779 GTGTGATGAGACAGATTTTGGACATTAATGTAAAGGTTTCTACGGGGAATCATTTGGTG 1838
OY 800 tggtagatcagaagacttttggctcgagacacatcagtaacatcaaatgataatgac 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1839 TGGTAGATGAGAACCTGTGATTTGTTGCCACACGATAGCTAATAAATTCCTGTATGGCC 1898
OY 860 gagatgagctgagctatgagaagatgagagcaagcaagggaaagcaaatgcatgtat 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1899 GTGAATAATCTCACATGATGATGATGAGAAAGCTGTCAAGAGCAATGCTATAGACT 1958
OY 920 ttatcatgagtttccataataattataatattgtaggggaaaaagagctcaaatga 979

```

```

DB 1959 TTATCATGAAACTGCTCATTAATTTGACACCCTGTTGGAAGAGAGGGGCCAGTTGA 2018
OY 980 gtgagggcagaaaagagagatcgcaattgtcgtgccttaagtcgaaaccccaagatc 1039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2019 GTGGTGGCGAAGAGAGAGATGCGCATTTGACAGTGCCCTGGTTGCAACCCCAAGATCC 2078
OY 1040 tgaatttaagtagcctgaagctgccttcgctgattcgaagcaagtaagctgttcaagctg 1099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2079 TCCCTGTGATGATAGGCGACAGCTGAGCCTTGAGACAGAAAGCAGACAGTTCAGGTGG 2138
OY 1100 cactggaagaagcgagcaaaagtgcggaataacgtgtgtagaacaacgacttctacta 1159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2139 CTCTGATTAAGCGCAAGAAAGTGGGACCCACCATTTGTGATGCTCTTTTGTACAG 2198
OY 1160 ttcgaaagtcaagattgtatgtgacccctaaagatggaatgctggtggaagaagagcac 1219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2199 TTCTGATGCTGACGTCATCGCTGTTTGATGATGAGAGTCTATTGTGGAGAAAGAAATC 2258
OY 1220 atgctgaactaatgcaaaaacagagctatattatcaactgtgatagtacaggaatata 1279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2259 ATGATGAACCTCATGAAGAGCAAGAGGCAATTTACTTCAACCTGTGCACAGACAGCAG 2318
OY 1280 aaaaagctgatacagatgagtlcaatgacatatcta-----ctga 1322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2319 GAATGATGATGGAATTAAGAAATGACAGCTGATGAATCCAAAGTGAATGATGCTTGG 2378
OY 1323 aagaagaaccaactactctctctgcactctgtggaagagataca----- 1366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2379 AAATGCTTCAAAATGATTCAGATCCAGTCTAATTAAGAAAGATCAACTCCTAGAGTGC 2438
OY 1367 -----agtcagacttcaatgacaagctgaggaatccaccacactaaagataagtc 1420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2439 TCCGTGATTCACAAAGCCCAACAGAGAAAGCTTAGTACCAAGAGGCTCTGATGAAAGTA 2498
OY 1421 ttctgaagctctctatataaaatttaagtaaaacagccctgaatgctgttgg 1480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2499 TACCTCCAGTTCCTTTTGGAGATTTGAAGCTTAAATTTAACTGAATGGCTTATTTTG 2558
OY 1481 ttctgggagacatgtgctcgtcttaaatggaactgttccatcttctccatct 1540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2559 TTGTTGGTATTTTGTGCCATTATTAATGAGAGGCTGCAACGACGATTTGCATTAATAT 2618
OY 1541 ttgcaaaatataacatgttggaaataatgataaaacacatlaag---catgat 1597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2619 TTTCAAGATATTAAGGGGTTTTCAGAAATGATGATGATCTCGAAGCAAAACGACAGATA 2678
OY 1598 cagaatattccatgatalctgcataatttggtgtattgtcttgcagttatctca 1657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2679 GTAACCTGTTTCACTATTTCTAGCCCTTGGAAATTTCTTTATTAACATTTTTC 2738
OY 1658 tgcaggaattatttaagcgagagagaggaatttaagatgagatlaagacacttg 1717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2739 TTCAGGGTTTCAATTTGGCAAAAGCTGAGAGATCTTACCAAGAGGCTCGGATACATAG 2798
OY 1718 ccttaaaagcattatatacagatattgcctggttggatgaaagaaagcaagcacag 1777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2799 TTTTCCGATCCATGCTCAGACAGAGATGTGAGTTGTTTCAATGACCTTAAAAACACACTG 2858
OY 1778 gagcttgacaacaatattagcatagatataagcaaatcaagaaagagcaacagttcca 1837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2859 GAGCATTTGCTACAGGCTCGCAATGATGTGCTCAAGTTAAAGGGCTATAGGTTCGA 2918
OY 1838 ggaattggcttcaacaataatgcaactaacaatgagacttcagttcatcttccatla 1897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2919 GCGTTCCTGCTAATTTCCCGCAATATAGCAAACTTTGGGACAGAGATATTAATATCTTCA 2978
OY 1898 tatatgagatgagatgacatccagatctgagatattgctccaglaacttgcgtgacag 1957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2979 TCTATGCTGGCACTAACAACGTGTTACTGTAGCAATTTACCAATCATTCATTAATAGCAG 3038
OY 1958 gaatgattgaaacccgagcaatgactgtattgtccaaagaataagcaagacttaagc 2017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 3039 GAGTGTGTAAGTAAATGTTTGGCTGGACAGCACTGAAAGATAAGAAAGAACTAGAAG 3098
Oy 2018 atgctggaaagataagcaactgaagcttggagaatatacgtactagtgatataacaa 2077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3099 GTGCTGGGAAGATCCCTACTGTGAAGCAATGAAACTTCCGAACCGGTGTTTGTGACATC 3158
Oy 2078 gggaaagacccctcagcaaatgtaagaagatgcttcaagatcaacaacaataacct 2137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3159 AGGACGACAGAGTTTGAACATATGTATGCTCAGAGTTTGAGGTACCATACAGAAACTCTT 3218
Oy 2138 cgaagaagcacagatatltggaagcgtltatgaltcaagccatgaccttatalatttg 2197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3219 TGAGGAAGACACACTCTTTTGGAAATTAATCTTCTTCAACCAGGCAATGATGATTTT 3278
Oy 2198 cctatgacagcgggttctgatttggagccttattatccaagctggagaaagaccag 2257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3279 CCTATGCTGGATGTTCCGGTGTGAGCCTACTGTGTGACATTAACCTCAAGACCTTTG 3338
Oy 2258 agggacgttcaatgcttcttactgcaatgcaatgcaatgcaatgcaatgcaatgca 2317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3339 AGGATGTTCTGTATATTTTCAAGCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3398
Oy 2318 tgcgttggctcctgcaataatcccaagccaaatcgggggcgtgcaatcgttgcctgt 2377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3399 GTTCAATTTGCTCTGACTATGCAAAAGCCAAATATCAGACGCCACATCATCATGATCA 3458
Oy 2378 tggaaagaacaaataatagacacgcgacgcaagaaggaaagcaagcaagcaatgtg 2437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3459 TTGAAAAAACCCCTTGTGATGACACTACAGCAGCAAGGCTTAATGCCGAACACATTTGG 3518
Oy 2438 aaggaatttagagttcgaagagctcttcttcatatcagtcgccaagtgttttca 2497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3519 AAGGAATGTCCATTTGGTGAAGTTGTATCACTATCCACCCGACCGGACATCCAG 3578
Oy 2488 tccctcgtggttaccctcagtalatgagcgaagaagacagtagcatgttggggaagca 2557
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3579 TGCTTCAGGGGACTGAGCTGGAGTGGAAGAGGCCAGACGCTGCTGTGGGGCAGCA 3638
Oy 2558 ggggcttggggaagaaagccttctgtcaactctgcagagacttattagccccgtgcaag 2617
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3639 GTGGCTGTGGAGAGACACAGTGTGCAGCTCTGAGCGGTTGACGACCCCTTGGCGAG 3698
Oy 2618 gacaagtgctgttctgtagtggtagtgaagaattgagtagtgcgttccg 2677
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3699 GGAAGTGTCTGCTTATGAGCAAGAAATRAAGCAGCTGATGTTCAGTGTGCTCCAGCAC 3758
Oy 2678 aataagcaatcgttccctcaagaagcctgtgcttcaactgagcaatgtctggaacatcg 2737
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3759 ACCTGGGATCGTGTCCAGAGGCCCATCTGTTGACTGACGACATTTGCTGAGAACATTTG 3818
Oy 2738 cctatgtagaacaagcgttggtagtgaatagtagatgaagaagccgaatgcaag 2797
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3819 CTTATGAGACACAGCGGGGTGTGTCCAGAGAGAGATCGTGAGGGGAGCAAAAGGAGG 3878
Oy 2798 caaatatcatcttcttatttgaaggtctccctgagaataacacacacacacacacacac 2857
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3879 CCAACATACATGCTTCAATCGAGTCACTGCGCTAATAAATATGACACTAAGTAGAGAGACA 3938
Oy 2858 aagagacagacttctcgtggcgagcaagaacaaagactagctatgcaagggctctctcc 2917
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3939 AAGGAACTCAGCTCTGTGGTGGCCAGAAAACAGCATTTGCTGCTGCTGCTGCTGCTG 3998
Oy 2918 aaaaacccaataatttatttatttatttatttatttatttatttatttatttatttattt 2977
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3999 GACACGCTCATATTTTGTCTTTTGGATGAGGCCAGCTCAAGCTGTGATACAGAAAGTGA 4058
Oy 2978 aggttggttcaagcatgccccttgaataagcgaagcgaagcgaagcgaagcgaagcgaagc 3037
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4059 AGTGTGTCACAAAGACCCCTGGAGCAAGACGAGAGAGGCGCACCTGTGATGATTTGCTGC 4118
Oy 3038 acagagctctgcaattcaagaacgagatttgaatggttctgcaacaatggaagataa 3097
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4119 ACCGCTGTCCACATTCAGAAATGACAGACTTAATAGTGTGTTGAGAAATGACAGAGTCA 4178

```

```

Oy 3098 aggaacaagacatcatcaagacccctcgtgagaaatcgagacatatatttaagttatga 3157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4179 AGGACATGTGGCAGCATGACAGCTGTGTGGCAGACAGAAAGCATATTTTCAATGTGCA 4238
Oy 3158 atgcacagtcag 3169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4239 GTGTCCAGGCTG 4250

RESULT
9
US-09-120-513-1
: Sequence 1, Application US/09120513
: Patent No. 6025160
:
: GENERAL INFORMATION:
: APPLICANT: Brun, Kimberly
: APPLICANT: Chenery, Richard
: APPLICANT: Ellens, Harma
: APPLICANT: Field, John
: APPLICANT: Yue, Lin
:
: TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY:
: ZIP: 19406
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/120.513
: FILING DATE: 22-JUL-1998
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: King, William T
: REGISTRATION NUMBER: 30,954
: REFERENCE/DOCKET NUMBER: GP50008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5015
: TELEFAX: 610-270-5090
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4233 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: Genomic DNA
: US-09-120-513-1

Query Match 32.8%; Score 1040.6; DB 3; Length 4233;
Best Local Similarity 60.3%; Pred. No. 4.4e-293;
Matches 1818; Conservative 0; Mismatches 1154; Indels 45; Gaps 4;

```

QY 380 ttcgtttcttcttagtgaatccatagccatttgcatttgagccagccctcact 439  
1014 TTACCGCTCTCTCTCATTTTATTTGGGACCTTTCAGTATTTGGCATTTTGGCCCAACA 1073  
QY 440 tgaacccctgcgaatagcccgagagctgccttccatttccaggtatgataaga 499  
1074 TAGAAGCTTTGGCAATGCAAGAGGGGACGCTTGAATCTTCAAGATATGATGATG 1133  
QY 500 aaccagatagataactttccacagctgatatataacctgaatccatagaagaactg 559  
1134 AGCCAAACATGACAGCTTCTCAACCAAGGACACAAACAGACAGATTAATGGGAAT 1193  
QY 560 tgaatttaaaatgcttcttcaattatccaaagaaccatcatcaaatlctgaag 619  
1194 TGGAAATTTAAATGTTTACTTCAACTACCATACAGAAAGTGAAGTTTGAAGG 1253  
QY 620 gtctgaatccagaataagctctgagagagacagtcgcttgctgcgtccaaatgacgtg 679  
1254 GCCTCAACCTTGAAGGTGAAGGCGGCGAGAGGTTAGCCCTGGTTGGCAACAGTGGCTGTG 1313  
QY 680 ggaagagtaaggtatgcagctctgcagaggtatattgatacgsatgatgacttatca 739  
1314 GGAAGACACAACTGTCCAGCTGCTGAGAGGCTCTAGAGCCCATAGAGGGCGAGGTCA 1373  
QY 740 tggtagatgaagaatgcacagagctttaaatgtagcgcattatcgaagacatttgag 799  
1374 GTATCGAGCGACAGAGACATCGAGACATGATGAGGTATCTCGGGAATTCATTGGGG 1433  
QY 800 tggtagatgaagaatgcacagagctttaaatgtagcgcattatcgaagacatttgag 859  
1434 TGGTAGATGAGAGAACCGGTGCTGTTGGCAGACAGATTGGCGAATAATTCGCTGACC 1493  
QY 860 gaagatgtagcagatgaagaagaatgagagagagagagagagagagagagagagag 919  
1494 GAGAAACGTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1553  
QY 920 ttaatcagagcttctcctaataataataatcattgtagagagagagagagagagagag 979  
1554 TCATCATGAACCTGCCACAAATTTAAACACCTGTTGGTAGAGAGGGCGACGCTGA 1613  
QY 980 gtggaagagcagaagaagaagaatgtagcgtgctgctgctgctgctgctgctgctgct 1039  
1614 GTGGGGGACAAACAGAGATGCCATTCGCCGGGCTGCTGCCGAAACCCCAAGATCC 1673  
QY 1040 tgattttagagagagcgtcgtccctgagattcagaagaagctcagctctcgaagctg 1099  
1674 TTTTGTGTGATGAGGCCACGTCAGGCTTGGACACAGAAAGCGAAGCCGTGTTCAAGCCG 1733  
QY 1100 cactggaagagcagagcaaaagctgcgaactacaatcgtgtagcacaacgacttttacta 1159  
1734 CTCGTGATTAAGGCTAAGAAAGGCGGACACCATGATGATGATGATGATGATGATGAT 1793  
QY 1160 ttcgaatgtagaatttgcattgtagccttaagaagatgtagcgtgcggaagaagaagac 1219  
1794 TGGCCAAATGCTGAGCTCATTTCTGTTTGTATGGGTGTATTTGTGATGAGGAAATTC 1853  
QY 1220 atgctgaactatgcaaaaagaaggtatattatcactctgtagatgtagcagaaggtat 1279  
1854 ATGAAGACCTCATGAAAGAGAGAGGCAATTTACTTCAAACTGTGTCATGACACGACTAAG 1913  
QY 1280 aaaaagcagatgaa-caagatgagatcaatgacatatcttactgaagaagaagaaccactca 1338  
1914 GAATGAATTAATGAACAGGAATATATGCTTATGATATCCCAAGTACACTGTGCTGTG 1973  
QY 1339 ctctcctgcacactgtgtagagacatcagactcagactcattgacaagagctgtagaactcc 1398  
1974 AGTTGACTTTCAGAAATAATCAAAATCTCCTTTAATGAAGAGATCAATTTGCGAAGATATCC 2033  
QY 1399 acccaatctaa-----agagataagctctc 1423  
2034 ACAGAAAGCAAGACCGAGAGAAAGACTTACTTGAAGAGAGAGTGTGATGAAATGTGC 2093

QY 1424 ctgaagctctctctataaaatlttaaglttaacaagcctgaatgagcctttgtgttc 1483  
2094 CTATGTTTCTCTTGTGGCAATATCTTAAGCTAAATATTATGATGATGATGATGATGATGAT 2153  
QY 1484 tggggaatctggcttcgtcttaaatggaactgtlccatcagaatcttccatccactcttg 1543  
2154 TGGGTACTTGTGGCTTATTAATGAGTGATACACCAAGTGTGTGCTATGATGATGATGATGAT 2213  
QY 1544 caaaatataacacagctt---ggaataatgataaaacacattaaagcatgagcaag 1600  
2214 CAAGATTTGAGGGGTTTTTCAAGAGACGACGACCATGAACCAACCAAGCAAGCAAGTGTGA 2273  
QY 1601 aaattatccatgatatctgcatttgggtgttattgtcttgcttgtaatttccatcag 1660  
2274 ACTTGTATCCCTTCTCTCTGCTGATGAGAAATGATTTCTTTGTTACCTACTTCTTTC 2333  
QY 1661 agggattatttccag 1720  
2334 AAGGCTTCACATTTGGCAAGAGCTGAGAGATCTCTACCAAGCGACTCCGATACATGATGCT 2393  
QY 1721 tcaagcagatgtatatacagagatattgcgtgttgaatgaagaagaagaagagagag 1780  
2394 TCAATTCATGCTCGACAGAGATTAAGCTGTTTGAATGACATTAATAACACCACTGGCT 2453  
QY 1781 gcttgacaacaataatgagcatagatatagcacaatccaagaagaagaagagagag 1840  
2454 CGCTGACTACAGGCTGCTAGTACGCTCTTAATGTTAAAGGGGCTATGGGCTCCAGGC 2513  
QY 1841 ttggcgtcttaacacaaatgagacatacagatgagacttcaagttatcatttccattat 1900  
2514 TTGCTGTAGTTACCCAAAGATGACAAACCTGGCACAGAAATTAATTTATTCCTTAATCT 2573  
QY 1901 atgagtgagagatgacatctcgtatctgtagtatctgctcagatctgcgtgtgacagag 1960  
2574 ATGGCTGGACACTTACACTTTTACTTGTAGTAATTAATTAATTAATTAATTAATTAATTAAT 2633  
QY 1961 tgattgaaacccgagcaatgagctgtagatttgcacaagaagaagaagaagaagaag 2020  
2634 TTATTTGAATGAATGAATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2693  
QY 2021 ctggaagaatgagacagcagagccttggagagatagactgactatagatgctatgaagag 2080  
2694 CTGGGAAGATGCTACAGAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2753  
QY 2081 aaaaagccttcagagcaaatgtagaagagatgtagcagcagcagcagcagcagcagc 2140  
2754 AGCAGAGTTTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2813  
QY 2141 agaaagcagagatattggaagcgtgtatgcatcagacagcagcagcagcagcagcagc 2200  
2814 AGAAAGCACACGCTTTTGGGATCACCTTCGCTTACCCAGGCGCATGATTAATTTTCT 2873  
QY 2201 atggaagaggttgcatttggagagcctattatcaagcgtgagagagagagagagag 2260  
2874 ATGCTGCTTGTTCGGTTCCGTTCCGTTCCGTTCCGTTCCGTTCCGTTCCGTTCCGTTCCGTT 2933  
QY 2261 gcatgtcatagcttcttactgcaattgcatatggaatggaatggaatggaatggaatgga 2320  
2934 ATGTTATGTTGATTTTCTCTGCTGTTGCTTGGTGCATGCGACGAGGAATACAGTT 2993  
QY 2321 ttgtgctcctgaaatccaaagcgaatcgaggagctgcagactgtgtgctgtgtg 2380  
2994 CATTCGCTCTGACTACGCGAAGGCCAAAGCTCAGAGATCCCATCATATCAGATTCATTTG 3053  
QY 2381 aaaaagaccaaataatgagagcgcgagctagaagaagaagaagaagaagaagaagaaga 2440  
3054 AGAAATATCCCGAGATGACAGCTACACAGAGAGGCTTGAACCTTAATTTGTTAGAG 3113  
QY 2441 ggaatttagcttcgagaagctcttctctctatcagctgcgcagagatgttctcactcc 2500  
3114 GAATGTGAATTTAATGAGATCATGTTCAACTTCCACCGACCAACATATCCAGTGC 3173  
QY 2501 tccgtggttatccctcagtatgagcgaagaagaagaagaagaagaagaagaagaagaaga 2560



Db 3174 TTCCAGGACTGAGCTTCAGAGGGAAGGCGCAACGCTTCGCTGGTGGGACACTG 3233  
Oy 2561 gctgtggaagaagcactctcttcaactctgcagagacttataaacccgtgcgaagac 2620  
Db 3324 GCTGGGGAAGAGTACATGTGGCTCAGCTGCTGAGCGCTTACAAACCCATGCTGGAA 3293  
Oy 2621 aagtcctgttgatgtgtgtatgtaagaagatgaatgtaacagtgagtgcttcgctccaa 2680  
Db 3294 CAGTCTTCTGATGAGCAAGAAATAAACACTCAACGCTCCAGTGGCTCGCGCCACC 3353  
Oy 2661 tagcaactcttccaaagcctgtgccttcaactcagactgtcagacatctgcagacatcgct 2740  
Db 3354 TGGGCAATTGTGCCAGAGGCCATCCTGTTGACTGCAGATCACCAGAAACATCGGCT 3413  
Oy 2741 atgtgtaaacaagccgtgtgtgcattagatgataaagaagccgcaaatgcagaa 2800  
Db 3414 ACGGAGAGCAAGCGCTGCTGCTCATGAGAGATGCTGAGGCGCCGCAAGGAGGCA 3473  
Oy 2801 atatcactcttattgaaggtctccctgagaaatagacaacagatgtgactgaaag 2860  
Db 3474 ACATCCACCACTTCATGCACTCACTGCTGAGAAATACACACAGAGTGGAGACAAAG 3533  
Oy 2861 gagcagacttcttgcgagcagaacaagaactagctattgcgaaggtcttcccaa 2920  
Db 3534 GGACTCACCTGTGGGGGGGAGAGAGCGCATGCGCATGCGGCGCCCTCTGTACAG 3593  
Oy 2921 aaccacaatttattgttgatgagcgaactcagccctcgataatgcagtgagaag 2980  
Db 3594 AGCTCACATCTTACTTGTGATGAGAGACATAGCTGTGATGAGAGTGAAGTAAAG 3653  
Oy 2981 tggttcagactgacctatataagcagagcaggaagacatgcttattgttcactca 3040  
Db 3654 TCGTCCAGAGAGCGCTGCAAAAGCCAGGGAAGGCCGCACTGCAATGTGATGCGGCA 3713  
Oy 3041 ggccttcgcaattcagaacgcagatttgatgtgtctgcacaaatggaagaatcaag 3100  
Db 3714 GCGTGTCCACCATCCAGAACGAGACTGTGATGCTGATTCAGAGAGCGCCAGTCAAG 3773  
Oy 3101 aacaagaactcattcaagagctcctgcagaaatgcagacatatatttaagttagtga 3160  
Db 3774 AGCAGGAGACCCACGACGAGCTGTGCGCCAGAAAGCATATTTCTGATGTTCAAG 3833  
Oy 3161 ccaagtcagtcagtgga 3177  
Db 3834 CTGGAGCAAGCGCTCA 3850

RESULT 10  
US-09-450-105-1  
: Sequence 1, Application US/09450105  
: Patent No. 6169166  
: GENERAL INFORMATION:  
: APPLICANT: Kimberly Anne Brun  
: APPLICANT: Richard James Chenery  
: APPLICANT: Harna Ellens  
: APPLICANT: John Anthony Field  
: TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
: FILE REFERENCE: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF  
: CURRENT APPLICATION NUMBER: US/09/450,105  
: EARLIER FILING DATE: 1998-07-22  
: NUMBER OF SEQ ID NOS: 2  
: SOFTWARE: FastSeq for Windows Version 3.0  
: SEQ ID NO 1  
: LENGTH: 4233  
: TYPE: DNA  
: ORGANISM: HOMO SAPIENS  
US-09-450-105-1

Query Match 32.8%; Score 1040.6; DB 4; Length 4233;  
Best Local Similarity 60.3%; Freq. No. 4,4e-293;  
Matches 1818; Conservative 0; Mismatches 1154; Indels 45; Gaps 4;  
Oy 200 ttgttaaggtatcacagatctcacaagatgcacaagatttggcattaaagaagactatag 259  
Db 840 ttgaaggtatcacataaatttagaagaagctcaaaagtgctgactaaagaagcattca 899  
Oy 260 ctccaagtgctctgtgtgtgttacttcttataatgaatgaatgaactgtgctt 319  
Db 900 cggccaacattccataglatgtcctacgtgtgtctatgtcgtctatgtacgtcgtcat 959  
Oy 320 ttgttgtagaagcctctgtattcttaattgagaacccggtatataccctgtgactgttc 379  
Db 960 tctgttatggagacccctctgtgctcct-ctcaaaagaaattctatgtgacaagtgc 1013  
Oy 380 ttgctgttctctttagtgaatccatagcaatgattatgtcatgtgagcaagctccctact 439  
Db 1014 ttaccgtctctctctatttatttgggacttcaagatgtgacattttagcccaaa 1073  
Oy 440 ttgaaccttcgcaatagcccgagagctgcttcaatttccaagttatgataaga 499  
Db 1074 tagaagccttgcgaatgcaagagggcgagcctatgaatcttcaagataatgttaag 1133  
Oy 500 aaccagatagatgaacttctccagactgtgataaaacccggaatccatagaaggaactg 559  
Db 1134 agccaagattgacagcttctccaagaagggacacaaacccagacagatataatgtgaaat 1193  
Oy 560 tggaaattcaaaatgttcttcaattcaatccacaagccatctcaagattcgaag 619  
Db 1194 tggaaattcaaaatgttcttcaattcaatccacaagccatctcaagattcgaag 1253  
Oy 620 gtctgaatctcaagattgaatgtcgaagaacagtcgctgtgtcgtcctcaatgagcag 679  
Db 1254 gccccaactgaaggtggaagagggcgagcagctgagccctgtgtgcacaagctgtgtg 1313  
Oy 680 ggaagagtagcgtgtgtccagcttctgcagaggtatataagatccgaatgagtattaca 739  
Db 1314 ggaagaagcacaactgtccagctgtcgcagaggtcttcaagaccctcaagagcgaggtca 1373  
Oy 740 tggtagtagagatcacatcagagcttcaaatgttcgagcatatcagaacatattgag 799  
Db 1374 gtatcgaagcagagcaactcgcagacatcagatgtatgtatctgtggaataatctggg 1433  
Oy 800 tggtagtcaagagcctgttctgttcggaaccacatcagatacaatcaatgagac 859  
Db 1434 tggtagtcaaggaaccgtgtgtgttcgacacagattgtccgaaacattcgtctagcc 1493  
Oy 860 gagatgattgtagtgaatgaagatgagagagagcagaaggaagcaaatgcgtatgatt 919  
Db 1494 gagaagaacgtccacatgagtagagagaagcgtcacaaggaagcgaatgctatgact 1553  
Oy 920 ttatcatgaggttccctataaattatacatgtgtagggggaaggaagagctcaatgga 979  
Db 1554 tcatcatgaactgtcccaacaatttaaacacctgtgttggtagagagggcgagcgca 1613  
Oy 980 gtgagggcagaacaagagatgcgaattgtcgtgcttgaatgcgaaccccaagattc 1039  
Db 1614 gtaggggagcagaagaagagatgcgaattgtccgggacctgtgtccgaaccccaagattc 1673  
Oy 1040 tgatttagatgaggtcgtgctgacctgagatgcgaagaagcgaatgctgttaagtg 1099  
Db 1674 ttgtgtgtagtggcccgctgaagccttgggaacagaagaagcgtgtgtgttcgagcg 1733  
Oy 1100 cacttgagaagcgagcaaaagtcgagactacatcgtgtgtagcaacagacttctacta 1159  
Db 1734 ctctgtagaaggttagaagaagcgagcagccacatgtgtatgtctcaacgctgtgtcaag 1793  
Oy 1160 ttcgaagtgcagatttgattgtgacctaaagatggaatgctgtgcggaagaagagcag 1219  
Db 1794 tgcgaatgtcagctcatctgtgttctgtatgtgtgtgtcatgtgtgagaagaagaaatc 1853







```

Qy 2340 caaagccaaatcgggctgcgcacatcgttcttctgttgaagaagaaccataataga 2399
    |||||
Db 1501 CAAAGCCAAATATCAGAGAGCCACATCATCATGATTCATTAAGAAACCCCTTTGATTGA 1560
Qy 2400 cagcgcagatcaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 2459
    |||||
Db 1561 CAGCTACAGCAGGAGAGGCGCTTAATGCGCAACATTTGGAAGAAATGTCACATTTGGTGA 1620
Qy 2460 agctctctctctctctctctctctctctctctctctctctctctctctctctctct 2519
    |||||
Db 1621 AGTTGTATTCACATATCCACCCGACCCGACATCCCACTGCTTACAGGATGACCTTGA 1680
Qy 2520 taticagcagagaaagacagatgacatctgttgggagcagcgcgttgggaaagacattc 2579
    |||||
Db 1681 GGTGAAGAGAGGCGCAGACGCTGCTGTGTGGGACAGAGTGGCTGTGGAGAGACACAGT 1740
Qy 2580 tgtcaactctgcagagacttataagaccctgtcagaagaagaagaagaagaagaagaaga 2639
    |||||
Db 1741 GGTCCAGCTCTGAGGCGGTTCTACGACCCCTTGGCAGGAAAGTGTGCTTGTGATGCGAA 1800
Qy 2640 ggaatgcaagaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 2699
    |||||
Db 1801 AGAATATTAAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Qy 2700 gctctgtctctcaactgcagcactgtctgagaaacatgcgctatgttgaacaacagccgtgt 2759
    |||||
Db 1861 GCCCATCCTGTTGACTGCACATGCTGAGAACATTCCTATGAGAACACACAGCCGGGT 1920
Qy 2760 ggtgcatatgataatgataaagaagccgcaaatgacgaataatccattcttatttga 2819
    |||||
Db 1921 GGTGTACAGGAGAGAGATGATGAGGCGCAGCAAGAGAGCCCAATACATGCTTTCATCGA 1980
Qy 2820 agtctccctgaagaatacaacaacaagaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 2879
    |||||
Db 1981 GTCACCTGCTATTAATATATACACTAAAGTAGAGACAAAGAACTGACCTCTGTGGG 2040
Qy 2880 ccagaacaagaactagactatgcaagggctctctcccaaaacccaatatttattgtt 2939
    |||||
Db 2041 CCAGAACAAAGCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2940 ggaatgagccactcagccctcagataagacagatgagaaggtgttgaagtcagtccttga 2999
    |||||
Db 2101 GGATGAAGCCAGTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
Qy 3000 taaagccagagcagaaagaaatgcctagctgctcactcagcagcgtctctgcaatcagaa 3059
    |||||
Db 2161 CAAAGCCAGAGAGGCGCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Qy 3060 cgcagattgataatgctgtctgcacaatgaaagataaagaagaagaagaagaagaagaaga 3119
    |||||
Db 2221 TGCAGACTTAATAGTGTGTTTCAGAAATGCGAGATCAAGATGATGATGATGATGATGATG 2280
Qy 3120 gctcctggaatcgagacatatatttaagttagtgaatgacagctag 3169
    |||||
Db 2281 GCTGCTGCGACAGAAAGCATCTATTTTTCATAGTGTGCTGACGCTG 2330

```

## RESULT 12

US-08-996-545-1

Sequence 1, Application US/08996545

Patent No. 5928898

GENERAL INFORMATION:

APPLICANT: Skatrud, Paul L.

APPLICANT: de Waard, Maarten A.

APPLICANT: Peery, Robert B.

APPLICANT: Andrade, Alan C.

TITLE OF INVENTION: Multiple Drug Resistance Gene atld of

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

```

STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,545
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4002 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4002
US-08-996-545-1

```

Query Match 13.4%; Score 426.6; DB 2; Length 4002;

Best Local Similarity 48.3%; Pred. NO. 3.9e-114;

Matches 1467; Conservative 0; Mismatches 1469; Indels 99; Gaps 6;

```

Qy 224 aagatgcaagaatcttggcataaagaagactatagcttcaaaagtctctgtgtgtgt 283
    |||||
Db 956 ACGAAGCTGAGAAATGGGGAAACAAAGAACAGATTGTTCATGGGTTTCATGATGGCCCA 1015
Qy 284 tgaactcttataatgaaatgaaactatggaactgtcttctgttgaatgaaactcttgaatc 343
    |||||
Db 1016 TGTGTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1075
Qy 344 ttaatgagaacccgtgataaccatcggaactgtcttctgttcttgaatgaaactcttga 403
    |||||
Db 1076 TAGATGG-----TGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1129
Qy 404 atagcaatgattgacatggaagcagatccctcacttgaacccctgcgaatgacccag 463
    |||||
Db 1130 TCGGATGCTTCTCTGTTGGGAAAGCTTAGTCCAAATGCTCAAGATTTTACAAAGCTGTGG 1189
Qy 464 gactgtccttcatatttccaggtatgataaagaaccagatagaaacttcca 523
    |||||
Db 1190 CCGCGCCGCAAAAGATATTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1249
Qy 524 cagctgatatataaaccctgaaatccatagaagaactgtggaattaaatgattcttcca 583
    |||||
Db 1250 ACGAAGGGAAGACCTGACCATTTTGAAGGCGCAATTTGAATTCGCAAAATGTCAGACATA 1309
Qy 584 attatcaatcaagacatctatcaagaattcgaaggctcgaatcgaatcaagaattagctg 643
    |||||
Db 1310 TTTACCATCTAGACCCGAGGTACCGGTCAATGAGAGATGTTTCTGTCAATGCCCTG 1369
Qy 644 gagaagagctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 703
    |||||
Db 1370 GAAAAACAAACCGCTTATGATGCGGCCCTGTGCTGTGAAAAAGATACGCTGTGCGCTTGG 1429
Qy 704 tgcagagttatagatcccgatgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 763
    |||||
Db 1430 TTGAGCGATTCTACATGCTGTTGCGGATACGCTTTTGTGATGAGCCATGATCAAGG 1489

```

Oy	764	ctttaaattgfcgpgcatatcagagaccatattggaatgtaagtaagacctgtttgt	823
Db	1490	ACCTTAATCTCCGCTGGCTTCGGCAACAGATCTTTGGTTAGCCAGAGCCCTGTCTTT	1549
Oy	824	tcggagccacatcatgaatacatat-----caagta	856
Db	1550	TTGGCAGCAGCATTTATTAATAATTTAGCCAGCGTCTCATCGGCACAAATACGAAATG	1609
Oy	857	gacgagatgatlgaactgaatgaagagatggaagacgaaggaaacaaatgcat	916
Db	1610	AATCCGAGAGATTAAGTCCGGGAACTCATCGAGAACCGGCACAAAATGGGAATGCTCATG	1669
Oy	917	attatcatggaagtcttccataaatttaatacatatgtgaaggaaaaagagctcaa	976
Db	1670	ACTTTATTACTGCTTCCTTCCTTGAAGTTATGAGCCAAATTTGGGACGCGTGCTTTC	1729
Oy	977	tgaatggaaggcagaacaagaagatcgcaatgctgcgttcctgaatctgcaaccaca	1036
Db	1730	TTTTCAGGTGGCGCAGAAACACGCGCATTTGCAATGGCCCGCTGGTTATGATGCCAAAA	1789
Oy	1037	ttctgattttaagatgaggtcaagtcgtcccttgatctcagaagaagtcagctgttaag	1096
Db	1790	TCCGCTCTCCGAGTGAAGCTACTTTCGCGCTTGGACACAAAATCCCAAGCGCGTTCACAG	1849
Oy	1097	ctgcactggaagaagcgaagaagaagtcggaatacaatctgtgtgaacaacgaactcta	1156
Db	1850	CAGCTTTGGAGAGGGCAGCTGMAAGGCCGAATCTATTGTATTCGCTCATCGCTTTTCA	1909
Oy	1157	ctatctgaagtgcagatttgatgtgaacctaaagatgaaatgcatgctgcgagaaagag	1216
Db	1910	CGATCAAAAGGGCCACACACATCTTGTGGTTCTGCTCAATGGCAAAATTCGTGCACAGGAA	1969
Oy	1217	cacatgctgaactaagtcgcaaaacggggtctatattcatcttgatgtaacagata	1276
Db	1970	CTCAGCATGAATTTGGTTGACCCGGAGGGCTTATCGCAAACTTGGAGGCTCAACGTA	2029
Oy	1277	ttaaaaa-----gtatgaacagatgagat	1303
Db	2030	TCAATGAACAGAAAGAGTGAAGCGCTTGGAGAGCCGACGCGTGAAGTCTACAGATG	2089
Oy	1304	caatgacatatctactcgaaaagaagaccactactctctctgcactctgtgaagaca	1363
Db	2090	CAGATATTGGCCAAATATCAAAATCGCTGCAAGGCCATCATCCATCTCGACGGAAAAACCA	2149
Oy	1364	tcaagtcagactcatctgacaagctgag---gaatcaccccaactaaagataagtc	1420
Db	2150	CAACCATTTGAGCCGACGGGACCCACAAAGTCTTTCACAGCGCGATCTTCTTAANAAC	2209
Oy	1421	ttctctgaagtcctctataaaaaattlaag-----tta	1456
Db	2210	CCCCGAAACAACTCCCAAAATACTCATTTATGACGCTGCCTCAAAATTTGTGCTCTTCA	2269
Oy	1457	acaaagctggaatggccttttggtctciggaacatctgtctctctctctaaatgtaact	1516
Db	2270	ACCGCCTCGAATATCCCTACATGCTCATCGTGCTGTCTTCTCAAGTTAGCTGTGGTG	2329
Oy	1517	ttcatcagatatttccatcatctttgcaaaaaattlaacca-----tglttgsaata	1570
Db	2330	GCCAAACCCAGCAAGCAGTGTCTTATGCTTAAAGCCATGACACATCTCGGTCCCAAGAT	2389
Oy	1571	atgataaacacacattaagacatgacgagaaattatccatgatattgctcatlttgg	1630
Db	2390	CACATATTAGCAAGCTTCGACATGATGCGGATTTTCGGCATATGATGTTCTTCGAGTGG	2449
Oy	1631	gtgtattattgctctgcagtatattcatcgaaggaattatttaacgacgaagcagggaaa	1690
Db	2450	GTAATCATTCAGTTTATACGCAAGTCAACAAATGTCGTGCATTTGCCGATCTCCGAGA	2509
Oy	1651	ttttaacgatgagataagaacttggccttcaaaagcatgtaatacagaatattgct	1750
Db	2510	GACTTATTTCGTGCGGAGCAAGCAGCTCTTTCGGAGATCTCGGTCAAGCATATTCGTT	2569

[illegible]

Db 3650 ACGGCTTTATACAGTGTGTCAGACAGGAGGACATGTTGTCGGCGCCAAAAGCAAC 3709  
QY 2891 gactagctcttcaagggtctcttccaaaacccaatttatgttgtagtgccca 2950  
Db 3710 GGTGTGCGCATTCGCCAGCCCTTCTTCGGATCCAAAATCCCTTCTTCGTATGAAGCGA 3769  
QY 2951 ctccagccctcagataatgacagtgaaaggtgttcagatgcctctgaataagccagga 3010  
Db 3770 CGTACGCCCTCGACCTCCAGTCAAGAAAGTCTCCAGCGCGCTTGGATGCCGCTGCC 3829  
QY 3011 cgggaagacagctagctagtgtaactcagagctctctgcaattcagaacgagattga 3070  
Db 3830 GAGGCCGGAACCAATCGCGCTTGCACACCGACTCAGCAGATTCAAAAGCGGAGCTTA 3889  
QY 3071 tagtgctcgcagaaatggaagataaagaaagaaactcaagactcactgagaga 3130  
Db 3890 TCTATGTTTTCGACCAAGCAAGATGCTGANAAGCGAAGCGACGCACTGCTCAGA 3949  
QY 3131 atcagacataatatttaagtagtgaatgcacag 3165  
Db 3950 AAAAGGCGGTACTGACGAGCTGTCACACTTGAG 3984

## RESULT 13

US-08-996-545-3

Sequence 3, Application US/08996545

Patent No. 5928898

GENERAL INFORMATION:

APPLICANT: Skatrud, Paul L.

APPLICANT: de Waard, Maarten A.

APPLICANT: Peery, Robert B.

APPLICANT: Andrade, Alan C.

TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESSES:

ADDRESS: Eli Lilly and Company

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996,545

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-11766

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

TELEFAX: 317-276-2763

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 4002 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: mRNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-996-545-3

Query Match 13.4%; Score 426.6; DB 2; Length 4002;  
Best Local Similarity 35.1%; Pred. No. 3.9e-114;

Matches 1065; Conservative 402; Mismatches 1469; Indels 99; Gaps 6;  
QY 224 aagatgcaaaagatttggcacaataaaggactatgcttcaaaagtcttgtagtgctg 283  
Db 956 ACGAAGCUCAGAAUUGGGAAACAAAGAACCAUUGUCUAGUGGUUUCUAGUAGUGCGCCA 1015  
QY 284 tgaactcttataatgaaacctaagactgtctctttttagttagtgaactcttgatcc 343  
Db 1016 UGUUUGGCCUUUAGUACUAGAACUACGUGUUGGUCUUGGUAUGUGGUUUCUGUUG 1075  
QY 344 ttaatggagaacctggtatataccatcgagacgtctctgtctgttcttcttagtgaacc 403  
Db 1076 UGAGUGG-----UGCAGUCGAGUGGUGGUAUUCUACAUUUCUAGUGGCCAUUCUGA 1129  
QY 404 atagcagttatgtcatcttgtagcagcagtcctcactcttgaacacttcgaatagccgag 463  
Db 1130 UCGGAUCGUCUCCUGGGGAAACGUUAGUCCAAUUGCUCAACACAUUUNCAAAAGCGUGG 1189  
QY 464 gagctgccttcataatltccaggtatgataaagaacccagtagataactttcca 523  
Db 1190 CCGGCGCCGCAAAAGAUUUGAGACGAGUCCGACAGUCCCAUUDAGUCCAUUAGUGA 1249  
QY 524 cagctgataataacctgaatccatagaagaagacgtggaatttaaatgtttcttcca 583  
Db 1250 ACGAAGGAGAGCGUCGACCAUUDUGAGGCGCACAUUGAGUACGCAUUGCAAGCAUA 1309  
QY 584 attatccatcaagacacatcatcaaatcttgaaggtctgactcgaatgaatgactg 643  
Db 1310 UUUACCCAUUAGACCCGAGGUCACCGUAGUGAGAGUUGUUCUUGUACUAGCCCGCUG 1369  
QY 644 gagaagacgtgccttgcgttgcgttctcaatgagcagtggaagatagcgttagtccagcttc 703  
Db 1370 GAAAAACAACCCCUUUAUGUGCGCCUUGGCUUGGAAAAAGUACGUGUGCGCUGG 1429  
QY 704 tgcagaagtataatgatccgagatgactgactatcaatgtagtgaatgaatgacatcagag 763  
Db 1430 UUGAGGGAUUCACAUUGGUCGCGGACGUGUUGUGUGAUGCGCAUUGCAUACAGG 1489  
QY 764 cttaaatgtgcgactatcgagacacatctgagtgtagtgtagtgaagacctgttctgt 823  
Db 1490 ACCUCAUUCUGCGUGGUCUGCCCAACAGAUUCUUGUUGUUGCCAGAGCCUGUUCUU 1549  
QY 824 tccggaccacatcagtaacaatat-----caagatg 856  
Db 1550 UUGGACGACGAGAUUUAAGAAUUAUAGCAGGUCUACUCCGCAAAAGUACGAGAAUG 1609  
QY 857 gacgagatgactgactgaatgaagagatggaagagcaggaaggaatgtgctatg 916  
Db 1610 AAUCCGAGGAUUAAGUUCGCGGAACUUCAGAGACGCGCAAAAAGGCAUUGCUCAUG 1669  
QY 917 attatactgagatttcctaataatlaacatctgtgtaggggaaaaagagctcaaa 976  
Db 1670 ACUUAUUAUACUCCUUGGUAAGUUAUGAAGCAUUGUGGCGACGCGUUCUUC 1729  
QY 977 tgaatggaggcagaagaacagagatgcgaattgctcgtgacttagtgcgaaccccaaga 1036  
Db 1730 UUUCAAGUGGCGCAAAACAGCCCAUUGCAUUGCGCCGUGCCGUUUAUGUACCCAAAAA 1789  
QY 1037 tctgatttaatgtaggtagctgctgccttgatctgaagaagcagtcagtgctcaag 1096  
Db 1790 UCCUGCUCUGGAUAGGUAUCUUGCGCCUUGGCAAAAUCCAGAGCGGUGGUCAAG 1849  
QY 1097 ctgcaetgagaagcagcagaagctgcagactacaatctgtgtagacacgacttcta 1156  
Db 1850 CAGCUUUGAGAGGCGCAGCUGUAGAGCCGCAACUACUUAUUGCUCAUUGCCUUCUUC 1909  
QY 1157 ctatcgaaagtgcagatttgatctgtgacccitaaagatggaatgctgctgcggaagaagag 1216  
Db 1910 CGAUCAAAAGCGCGCACACAUUUGUGGUUGGUCUAAUUGCAAAAUUGCUCAAAAGGA 1969  
QY 1217 cacatgctgaactaatgagcaaaaagaggtctatatattcacttgtagtgaagata 1276  
Db 1970 CUCACAGUAUUGUGUAGCCGCGAGGCGCUUUAUUGCAAAAUUGUGGAGGCUCAACGUA 2029

QY 1277 ttaaaaa-----gctatgaacagatgagct 1303  
Db 2030 UCAAGAAACAGAGAACGUCAGCCUUGAGAGACGCCGACCGUAGUCCUACAGAAUG 2089  
QY 1304 caatgacatatctctgaagaagaacacactcctctcctgactctgtgaagaga 1363  
Db 2090 CAGAAUUGGCCAAAACUAGUGGUCAGGCAUACCCAGUUCUCCAGGAAACCA 2149  
QY 1364 tcaatgcagactctatgaagaagctgag---gaatcacccacttaagaagatc 1420  
Db 2150 CAACCAUUGAGCCGACGCGACCCACAAAGUCUGUUCAGCGCAUUCUUAAGAGC 2209  
QY 1421 ttcctgaagctctctatataaatttaag-----ttaa 1456  
Db 2210 CCCCAGAAACAACUCCGAAAUACUUAUUGAGCGUCGUAUUAUUGUCCUCCUUA 2269  
QY 1457 acaagcctgaatgagctctgtgtctgagacatgtgctctgttctaagtgaactg 1516  
Db 2270 ACCGCCUGAAUUCGCGUACUAGCUCUAGCUGUUCUUCUACUUGUAGCUGGUG 2329  
QY 1517 tctacagatattcactcactcttgcaaaaataataca-----tgttggaata 1570  
Db 2330 GCCAACCCAGCAAGCAGUGCUUAUUGCUAAAGCAUACACUCCUCCGCGAGAAU 2389  
QY 1571 atgataaaacacataaagcagatgacgaatattatctcagatctgctatctg 1630  
Db 2390 CACAUAUAGCAAGCUCGACUAGUAGCGAUUUCGUAUUGUCCUUGGUGUG 2449  
QY 1631 gttatttctgtctgactgaattatctgacagagatttaagcagagagagga 1690  
Db 2450 GUUAUACUAGUUAUACAGCGUACACCAUAGUGUGCUGCAUUCUCCGUAUUCGAGA 2509  
QY 1691 ttttaacagatgaatgaacactgtgacctcaaaagcattatatacagattgct 1750  
Db 2510 GACUUAUUGUGCGCGAGAACACUCCUUCGAGCAUACUCCGUAAGCAUUGCU 2569  
QY 1751 gtttgaagaaagaaacagcacagagagcttgacaaatatagcctatagatag 1810  
Db 2570 UCUUUGACAAGGAAGAAUAGACCGCGGUCUAGACCUUUCUUGUCGAGAGACA 2629  
QY 1811 caacaatcaagagcaacagatctcagatgtgcttcaacacaaatgaactaaca 1870  
Db 2630 AGCAUUCUUCGCGUUAUAGCGUAGCUUAGCAGCAUUCUAGACUCCUCCAGACC 2689  
QY 1871 tggagcttgcattatcattccttataatgagtgagagatgaatcctgattctga 1930  
Db 2690 UAGGAGCGGCUUACUUAUUGCCUUGGCAUUGGUGGAAUUGCCUUAUUGUACU 2749  
QY 1931 gtaatgctccagctactgctgcgagacagatgattgaaccgacagatgattg 1990  
Db 2750 CGGUUGCGCGGUCUUCUGGCAUUGCGGUUUCACCAUUAUUGCAGCCAGUU 2809  
QY 1991 ccaacaagaataagaacttaagcactgagacagtagcaactgaagcttgaga 2050  
Db 2810 AAUACAGCUCUACAGUUGUUAUAGGGAUUCGAAACUUCUUGCCAGCGCUAACU 2869  
QY 2051 atatacgtactatagtgcttaataagaagaaagccttcgacaaatgataaga 2110  
Db 2870 CUUUCGCGACAGUUCGUAUUAACCCGGAAGGAGAUUCUGGAGAUUUAACU 2929  
QY 2111 tgcctcagaactaacaagaataactcgaagaagaagcattatgaaagctgattg 2170  
Db 2930 ACUUGAGCCACAGCAGGACCAUUCUUAUUCUUGUUGGUCUACUCCUUAUUAUG 2989  
QY 2171 catcagacatgcttataatttgcctgaagcagaggttgattgagactatt 2230  
Db 2990 CGUUCUGCAGGACUUGUUAUUCUUGCGUUGCGUUGGUAUGGAGGAGAC 3049  
QY 2231 taatcaagctgagcaatgacccagagagcagtgatttaatttcaatgactgact 2290  
Db 3050 UUCUUGGUCACGACGAGUAGAUUUAUUCGCUUUGUUGUUCUCCGAGAUUUCU 3109

QY 2291 atgagactatgcccagaaagcgtctgttgcctcctgaatatctcaagcccaat 2350  
Db 3110 UUGGUGCUCAUCCCGGGACACCGUUCUUAUUGCAACAGACUAGGCAAGGAGGA 3169  
QY 2351 cggggcctgcgcactctgttgccttctgaaagaaacaaataagacagccgagct 2410  
Db 3170 AUGCGCGCGCAUUAUUCGAGACAGUUCUACACCAAGACCAAAUUAUUAUGUGUG 3229  
QY 2411 aagaaggaaaaagcagacacatgtaagggaattgagtgatttcgagaagctcttct 2470  
Db 3230 AAGAGGCGAGAACGUCGAAACGUGUAGGUAUUGGAUUAUUGGAAACGUCACUUA 3289  
QY 2471 tctacatgctccagatgtttcactcctcagtggttaccctcagatattgagag 2530  
Db 3290 GAUACCCGACCCGCAAGAACCCUUGUCUGCGGCGUUGAGACCUAGACCGUAGCCUG 3349  
QY 2531 gaaagacagtagactgtgaggaagcagcgtgtgaggaagacactctgttcaactc 2590  
Db 3350 GACAUUAUUGUGCGUUAUUCGAGACCGGUGUGGCAAGUACACAUUGCAUUGC 3409  
QY 2591 tgcagacattatgacccgtgcaaggaagtgctgttggatggtgtgatacaag 2650  
Db 3410 UUGAGCGCUUUAACGUAUUGCGGUGUCCUUGUUGUAGGGAAGCAUUAAGUA 3469  
QY 2651 aattgaatgacagtgctccttcccaatgacatgcttccaaagagcctgtgctct 2710  
Db 3470 AACUUAUUAUACUCCUACCCAGCAGUUCUUCUACUUGUACGCCAGAGCGGACACUG 3529  
QY 2711 tcaactgcagactgttgaacacatgcctatgtagaacaagccgtgtgtgcatatg 2770  
Db 3530 ACCAGGCGACCAUACAGGAACAAUUAUUGUUAUUGCAUUAUAGUACGUAACCGAG 3589  
QY 2771 atgagatcaagaagcgcgaatgacgaataatcattcttatttgaagctcctg 2830  
Db 3590 AAUUCUUAUUAAGCUCUUGCAAGGAGCAUUAUUAUAGCUCUACUUAUUGUCUCCCG 3649  
QY 2831 agaataacacacaaagtgtgacaggaagagcagacttctgagcgcaagaacaa 2890  
Db 3650 AAGGCUUUAUUAUUAUUGUGGCAAGGAGGCAUUGUUGUGCGGCAAAAGCAAC 3709  
QY 2891 gactagctatgcaaggctctctcacaacaaacaaatattatgttgaagagca 2950  
Db 3710 GUGUGGCCAUUUGCCGAGACCCUUCUUGGGAUCCAAAUUCUUCUUGCAUAGGGA 3769  
QY 2951 ctcaagccctcgatatagacagtagaaggtgttcagacatgccttataagccaga 3010  
Db 3770 CUCAGCCUCCACUCCGACUACGAAAGGUCUCCAGCGGCUUGGAGUUGCCGUGCC 3829  
QY 3011 cgggaagacatgcttagtgctactcaagctctcgaattcgaagaagcagattga 3070  
Db 3830 GAGGCCGAAACCAUUCGCGUUGCAACCGACUACGACGAUUAUUAAGGCGGAGUUA 3889  
QY 3071 tagtgctcagcaatggaagataaaggaagaagacatcaatcaagctcctgaagaa 3130  
Db 3890 UCUUAUUGUUGCAACAGGCAAGAUUGUAGGAAGGGAACGACAGCAUUGUCCAGA 3949  
QY 3131 atcagacatatatttaagttagtgaatgacag 3165  
Db 3950 AAAAGGCGCGUACUACGAGCUGGCAACUUDGACG 3984

RESULT 14  
US-09-328-320-1  
; Sequence 1, Application US/09328320  
; Patent No. 6228615  
; GENERAL INFORMATION:  
; APPLICANT: Skarud, Paul L.  
; APPLICANT: de Waard, Maarten A.  
; APPLICANT: Peery, Robert B.  
; APPLICANT: Andrade, Alan C.  
; TITLE OF INVENTION: Multiple Drug Resistance Gene atcd of  
; TITLE OF INVENTION: Aspergillus nidulans  
; NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/328,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/996,545  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4002  
US-09-328-320-1

Query Match 13.4%; Score 426.6; DB 4; Length 4002;  
Best Local Similarity 48.3%; Pred. No. 3.9e-114;  
Matches 1467; Conservative 0; Mismatches 1469; Indels 99; Gaps 6;

QY 224 aagatgcaaaagatttggcctaaagaagacttcaaaagtgcctcttggtg 283  
DB 956 ACGAAGCTGAGAAATGGGGAAACAAAGACCAAGATTGTGATGGGTTTCATGATTGGCCCA 1015  
QY 284 tgaactcttataatgaatgaactatggaacttcttcttgatgaaacctctgattc 343  
DB 1016 TGTGGGCTTATGTACGCACTACGCTTGGCTTGGATGGGTTCTGCTGTTCCG 1075  
QY 344 ttaatggaagaccgtgataataaccatcggaactgttcttgcgttcttcttgaatgcc 403  
DB 1076 TAGATGG-----TGCACGTGATGTGGTGATATTCTCACAGTTCTCATGGCCACTTGA 1129  
QY 404 atagaagattatgcatltagcagcagagccctcaacttgaaccttgcgaatagccgag 463  
DB 1130 TCGGATCGTTCCTCTGGGGAAAGTTACTCCAAATGCTCAAGCATTTTCAAAAGCGTGG 1189  
QY 464 gagcgccttcaatatttccaggtatgtaagaagaaccaglaagataactttcca 523  
DB 1190 CCGGGCGCGCAAGATATTTGGAACGATCGCCAGTCCCATTTAGATCATATTGGA 1249  
QY 524 cagctgataataaccctgaatccatagagaagaactgtggaatttaaaatgttcttca 583  
DB 1250 ACGAAGGGAAGCGCTCCATTTTGGGGCCACATTGATTAGCAATGTCAAGCATTA 1309  
QY 584 attaccatcaagaccatcatcaagattctgaaggtctgaatccagaattgaatgctg 643  
DB 1310 TTATCCCATCTAGACCCGAGGTCACCGTCAATGAGGATGTTTCTGTCAATGCCCCGCTG 1369

QY 644 gagaagacgtcgcccttggtcgtctcaatgacgtggaagagtagtgatccagcttc 703  
DB 1370 GAAAAACACCGCTTATAGTGGCCCTCGGCTGTGAAAAAATGAGTGGTGGCTTGG 1429  
QY 704 tgcagaggttatatgactcgcatgcatgacttcaatggtgataagaaatgacatcagag 763  
DB 1430 TTGAGCGATTCTACATGCTGCTTGGCGGTACGGTTTGGTGGATGGCCATGACATCAAG 1489  
QY 764 cttaaatgctcgccatlatcgagaccatattgagtggtttagtcaagaccgttttgt 823  
DB 1490 ACCATCATCTCCGCTGCTTCCGCAACAGATCTTTGGTTAGCCGAGCCCTGTTCTTT 1549  
QY 824 tcgggaccacatcagtaaacatat-----aaagtatg 856  
DB 1550 TTGGCAGCAGATTATTAAGAAATTAAGACACGCTCTACGCGACAACTACAGATG 1609  
QY 857 gacgagatgattgacgagatgagagatgagagcagaaggaagcaatgctatg 916  
DB 1610 AATCCGAGGATAGGTCGGGAACTCATGCAAGACGGGCAAAATGCGAATGCTCATG 1669  
QY 917 attatcatgaggttccctaaatlaatacatlgtlagaaggaaagagctccaa 976  
DB 1670 ACTTATTAAGCTTCCCTGCTGAAGGTTATGACCAATGTTGGGCAAGCTGCTTCTCC 1729  
QY 977 tgaagtgaaggcagaagaagagatgcgaattgctgctgacttgaatccgaaccaga 1036  
DB 1730 TTTCAGGTGGCCAGAAACACGCGATTCATTCGCTGCTGTTGATGACCCCAAAA 1789  
QY 1037 ttctgattttagatgaggtcagctgccttgatctagaagaagaatcagctgttcaag 1096  
DB 1790 TCTGCTCTCTGATGAAGCTACTTCTGGCTTGGACACAAATTCGAAGGGGTGGTTAAG 1849  
QY 1097 ctgacgtggaagcagcagcagaaggtcgaactacaatcgtlgttagcacacagacttcta 1156  
DB 1850 CAGCTTTGAGAGAGGCGACGTGAAGCGGACACTACTATTGTGATCGCTCATCGCTTTCCA 1909  
QY 1157 ctattcgaatgacgaattgattgtgacctaaagagatgaatgctgctgaggaagaagag 1216  
DB 1910 CGATCAAAACCGCCACACACATTTGGTGTGCTGCTCAATGCGAAATGTGCAACACGAA 1969  
QY 1217 cacatgctgaactaattgcaaaacagagtcataattcaactgtgattcacaagata 1276  
DB 1970 CTCACGATGAATTTGTTGACCGCGGAGCGCTTATCCGAAATCTGTGAGGCTCAAGTA 2029  
QY 1277 ttaaaaa-----gctgatgaacaagatgagt 1303  
DB 2030 TCAATTAACACAGAAAGAGCTGACGCTTGAAGACGCGGACGCTGAGAGATCTCAGATG 2089  
QY 1304 caatgacatattctactgaaagaagaaccaactcaactctctgcatctgttgaagaga 1363  
DB 2090 CAGATATTGCCAAATCAAAACTCGCTCAGCGCATCATCCGATCTGACGGAACCA 2149  
QY 1364 tcaatgacacttcatgacagagctgag---gaatcacccaactcaagaagatgac 1420  
DB 2150 CACCATTTGACCGACGAGGACCCACAAAGTCTGTTTCCAGCGCATCTTTCTAAAGAC 2209  
QY 1421 ttccctgaagctctctatlaaaaatttbaag-----ttaa 1456  
DB 2210 CCCCCGAAACAACTCCGAAATATCATATGAGACGCTGCTCAAAATTTGTGCTTCTTCA 2269  
QY 1457 acaagcctgaatgaccttltgtgttcggygaacttggcttctgttctaaatgaaactg 1516  
DB 2270 ACCGCCCTGGAATCCCGTACATGCTCATCGCTTGTGCTTCAGTGTATCTAGCTGGTG 2329  
QY 1517 ttcatccaglatlcttcaatcactcttgaanaattataacca-----tgttggaaata 1570  
DB 2330 GCCAACCCACGACACAGCTGCTATATGCTTAAGCATATGACACTCTCGCTCCAGAT 2389  
QY 1571 atgataaacaacatlaaagaatgataagaatlaattccatgataacttctgtg 1630  
DB 2390 CACATATAGCAGAGCTTGACATGATGCGGATTTCTGTATGATGTTCTTCTGTTG 2449  
QY 1631 gtgtatttgccttltgcatgattatltcatgacaggaatttlaagcgacagcaggggaaa 1690



Db 2450 GTATCATTCAGTTTATCAGCGATCAACCAATGTCGATTCCTATGCTCCGAGA 2509  
Oy 1651 tttaacgaatgaagtaaacacttgccctcaaaagcaltatataagattgacct 1750  
Db 2510 GACTTATTCGTGCGCGAGCACTGCTTCGAGCATCTCCGACAGACATTTGCTT 2569  
Oy 1751 ggtttgagaaaagaaaacagacagagcttgacacatatatgaacatagatag 1810  
Db 2570 TCTTTGACAGAGAAAGAAATAGCACCGGCGCTGACCTCTTTCTGTCACCGAGAGA 2629  
Oy 1811 cacaatcaagagcaacaggttccagatgtgcgtcttaacacaaatgaacaa 1870  
Db 2630 AGCATCTCTCGGTTGTAGCGGTGTGACTTAGAGACATCTTGATACCTCCACAGACC 2689  
Oy 1871 tggagcttcagatcaatcccttatatatgataggagatgaacatccatctga 1930  
Db 2690 TAGGAGCGCTATCATTTATTCCTGCGCATTTGGGATAATTTGGCTTAGTTGTACT 2749  
Oy 1931 gtattgctcagacttgccgttgacaggaatgtgaacccagcaatgactgtattg 1990  
Db 2750 CGGTTGTCGGGTTCCTGCGATGCGGTTCTACCGATCTAATVGTAGCCAGTTTC 2809  
Oy 1991 ccaacaaagataagcaaacacttaagcaltgctgaaagataagcaactgaacttgaga 2050  
Db 2810 AATCAGCTCCAGGTTGCTTATGAGGATCTGCAAACTTTGCTTGACAGCTACATGCT 2869  
Oy 2051 atatacgtactaagtgatcaatcaacagaaagacacttgacaaatgtatagaaga 2110  
Db 2870 CTATTCGCGACAGTTGCGCTAATTAACCCGGAAGGATGTGGAGATTAACATGCC 2929  
Oy 2111 tgcctcagactcaacagaaataacctcgaaagaaagcaatatttgaaagctgtatg 2170  
Db 2930 AGCTTGACGCACAGAGCGACGACACTTAATCTCTGCTTGAGCTACCTGTTAATAG 2989  
Oy 2171 catcagcactgacttataatatttgctcactgacagaggttctgacttgagacctatt 2230  
Db 2990 CGTGTCCGAGGCACTGTTTCTTCGCGTTGCGCTGGGTTTGGTACGAGGAGACAC 3049  
Oy 2231 taatcaagactgagcaatgaacccagagagcatgtatcatagttttctgcaattgat 2290  
Db 3050 TTCTTGGTCACACAGATGATGACATTTTCCTCTCTTGTGTTCTCCGAGATTCCT 3109  
Oy 2291 atgagactatgccaatcgaaagaaagcgtcttgctcctcgaatattccaaagccaat 2350  
Db 3110 TTGGTGTCAATCCGCGGACCGCTTCTTTCCTTTCACACGACATGCGCAAGGGAAGA 3169  
Oy 2351 cggagggctgcatctgttctgctgttgaaagaaacaaatataagacgcccagctc 2410  
Db 3170 ATGCGCGCGCGAATTCGAGAGACTGTGACCGAAGCCACAAATGATATGATGCTGCTG 3229  
Oy 2411 aagaaggaagaaagcagaacatgttgaaaggaattagatttcgagaagctcttct 2470  
Db 3230 AAGAGGGGAGAAAGCTCGAAACGAGTGAATGAAATTTAGGAACGCGACTTCA 3289  
Oy 2471 tctatcaatgcccagatgttttcaatccctcoggggtatccctcagttatgagag 2530  
Db 3290 GATACCCGACCGCCCAACACAGCTGTCTGCGGCGCTTGACCTGACCGGGAAGCTCG 3349  
Oy 2531 gaaagacagtagcatlttgaggagcaagcgtgaggaaagcaactctgttcaacttc 2590  
Db 3350 GACATATATGTTGCGGTTGTGCGAGCCGCTGTGCAAGAGTACCAACATGATGATGC 3409  
Oy 2591 tgcagagacttaagaccggtgcaaggaagtgctgttgatggtgtgagtgcaaaag 2650  
Db 3410 TTGACCGCTTTACGATCGATTGCGGGTCCATCTTGTGATGGGAAGGACATAACTA 3469  
Oy 2651 aatgaatgaatgagtcctcgttcccaatagcaatgcttccctaagaagcgtgacct 2710  
Db 3470 AACTAATATCAACTCTACCGCAGCTTCTGTCACTGTACGCGAGAGCGCACATGCT 3529  
Oy 2711 tcaactgacatgtctagacaatcgcctatgtgacacaagccgtgtgtgacattag 2770

Db 3530 ACCAGGCGACCATCAAGGAAAAACATCTACTTGGTATGTGGAAGATGACCTACCGAAG 3589  
Oy 2771 atgagatcaagaagcgcgaatcgacgaataatcattcttaattgaaggtctccctg 2830  
Db 3590 AATCTTATTAAGGCTTGCAAGGACGCTAATATCTACGATCTCATATGTCGCTCCGG 3649  
Oy 2831 agaaataacacaagaatgtgactgtaagagacacagcttcttgaggccagaacaa 2890  
Db 3650 AGGCTTTAATACATCTTTTGGCAGCAAGGAGGAGATTTGTCTGGCGCCAAAAGCAAC 3709  
Oy 2891 gactagctatgcaagggctcttctccaaaaacaaatttaltgttgatgagcca 2950  
Db 3710 GTGTGGCATCTGCCGAGCCCTTCTTCGGGATGCCAAATCCTTCTTCGATGAAGGA 3769  
Oy 2951 cttaagcccttgataatgacagtgagaagtggttcaagatgacctgtataagccaaga 3010  
Db 3770 CGTCAGCCCTGACATCCAGTACGACGAAAGGTCGTCACAGGCGCTTTGATCCGCTGCC 3829  
Oy 3011 cgggaagagacatgacctagtgctcaactcaaggtctctgcaatcagaagcagaattga 3070  
Db 3830 GAGGCCGAACCAATCCGCTTGACACCGACTACGACGATTCAAAAGCGGACGTTA 3889  
Oy 3071 tagtgttctgcacaaatgaaagataaagaaagcaagaaactcaatcaagctctgagaa 3130  
Db 3890 TCTATGTTTTCGACCAAGCAAGATGTCGAAGCGGAACGACGCACTGCTCAGA 3949  
Oy 3131 atcgagacatatatttaagttagtgaatgcacag 3165  
Db 3950 AAAAGGCGCGGTACTACGAGCTGTGCAACTTGACAG 3984

RESULT 15  
US-09-328-320-3  
Sequence 3, Application US/09328320  
Patent No. 6228615  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Maard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atld of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/328,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/996,545  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid





Job time: 12672 sec

```
Db 2930 AGCUGAGCCACAAAGCAGACACUUAUUCUUCUGAGUCUAUCCUUAUAUAG 2989
Oy 2171 caticagccatgaccttatatttgctctatgacagaggttcgattgagcctatt 2230
Db 2990 CGUCGUCGACAGCACUUAUUCUUCGUCGUCGUCGUCGUCGUCGUCGUCGUCG 3049
Oy 2231 taatcaagctggaagaaatgacccagagagcatgttcatagttttactgcaatgcat 2290
Db 3050 UUCUUGGUCACACAGAUUAUAGCAUUAUUCGCUUUAUUAUUAUUAUUAUUAUUAU 3109
Oy 2291 atgagagctatgccaatcggaataacgctcgttttgctcctgaaatattccaaagccaaat 2350
Db 3110 UUGGUGUCUAUUCGCGGCGACCCGUCUUAUUCUUGCAGACAGCAUAGGCAAGCGAAGA 3169
Oy 2351 cgggggctgcgcacatctgttctcctgttggaataaaaccaaataatagacagcgacatc 2410
Db 3170 AUGGCGCGCGCAAUUCCGACGACGUCGUCGACCGAAAGCCACAAUUAUUAUUAUUAU 3229
Oy 2411 aagaagggaaagacacacacatgtgaaaggaattagaagtttcagagaagtccttct 2470
Db 3230 AAGAGGCGGAGAGAGCAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3289
Oy 2471 tctatccatgctgcccagatgttctcctcctcgtgcttaccctcagttatgacgag 2530
Db 3290 GAUACCCGACCCCGAGACAGCCUUGUCUUGCGCGCUUUGGACCUAGACCGUAGGCCUG 3349
Oy 2531 gaaagacagtgatcttggtgggagcagcgctgtgggaaagacatctctgttcaactc 2590
Db 3350 GACAAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3409
Oy 2591 tgcagaagacttatgaccctgcaagacaaagtgtgttgatgtgtgagtgcaag 2650
Db 3410 UUGAGCGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3469
Oy 2651 aatgaaatgtacagtgctcgttcccaataagcaatgcttcctcaagagccgtgtctct 2710
Db 3470 AACUAAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3529
Oy 2711 tcaactgacgatgtgtgagaacatcgctatgtgacaaacagccgtgtgtgacatag 2770
Db 3530 ACCAGGCGACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3589
Oy 2771 atgagatcaagaagccgaaatgacgaatcatcttcttatttgaagtcctcctg 2830
Db 3590 AAUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3649
Oy 2831 agaaatacaacacacaaatggaactgaaagagcagcttctgctgagccagaaacaa 2890
Db 3650 AGGCGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3709
Oy 2891 gactagctatgcaaggtctcttccaaaacccaaaatttatttatttatttatttattt 2950
Db 3710 GUGUGGCCAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3769
Oy 2951 cticagacctgataatgacagtgaggaagtggttcacatgcttctgataaagccagga 3010
Db 3770 CGUCAGCCCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3829
Oy 3011 cgggaagacatgctcctagtgctacacaggtctctgcaattcagaagcagattga 3070
Db 3830 GAGGCGGAACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3889
Oy 3071 tagtggtctgcacaaatggaataaaggaagaaactcataagagctcctgagaa 3130
Db 3890 UCUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3949
Oy 3131 atcgagacatatatttgaattgaattgaattgacag 3165
Db 3950 AAAAGGCGCGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3984
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:31:52 ; Search time 9392.76 Seconds  
(without alignments)  
3634.644 Million cell updates/sec

Title: US-09-873-409-12

Perfect score: 3177

Sequence: 1 atgctatcatcattgcacag.....atgcacagtcagtcagctga 3177

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapect 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hc:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	545	17.2	545	10	AL040762	AL040762 DKF2P434C
2	386.6	12.2	943	10	AL520322	AL520322 AL520322
3	385.8	12.1	405	10	AA243820	AA243820 2767406.6
4	357.4	11.2	559	11	BF692596	BF692596 602248949.4
5	298.8	9.4	894	11	BF584668	BF584668 60208406
6	275.2	8.7	726	11	BG293345	BG293345 602390738
7	273.2	8.6	1019	11	BG248052	BG248052 602359987
8	269	8.5	944	11	BF796582	BF796582 602258463
9	262.2	8.3	785	10	AV709991	AV709991 AV709991
10	253.2	8.0	780	11	BG587938	BG587938 EST489713
11	252.4	7.9	560	10	BE749379	BE749379 200194 MA
12	252	7.9	886	13	BH139685	BH139685 ENTN447TR

13	249.4	7.9	981	11	BE969667 602272046
14	249.2	7.8	932	13	A2670821 A2670821
15	249	7.8	795	11	BF313560 BF313560
16	248.4	7.8	880	13	A2687805 A2687805
17	241.4	7.6	886	13	A2540627 A2540627
18	239.6	7.5	871	13	A2682350 A2682350
19	239.6	7.5	897	13	A2541090 A2541090
20	238.8	7.5	853	13	A2679807 A2679807
21	237.2	7.5	834	13	A2548312 A2548312
22	234	7.4	947	13	A2687753 A2687753
23	232	7.3	503	13	AQ380214 AQ380214
24	230.2	7.2	823	13	A2532602 A2532602
25	229.4	7.2	482	11	B1344244 B1344244
26	227.8	7.2	651	10	BE016688 BE016688
27	226	7.1	687	11	BG646725 BG646725
28	225	7.1	919	13	A2690701 A2690701
29	224.6	7.1	891	13	A2682250 A2682250
30	224.2	7.1	842	13	A2671925 A2671925
31	219.8	6.9	661	11	BG302675 BG302675
32	215.2	6.8	899	13	A2548799 A2548799
33	214.4	6.7	921	13	A2687628 A2687628
34	213.8	6.7	575	10	AA342644 AA342644
35	213	6.7	519	10	AA434959 AA434959
36	212.2	6.7	605	10	AV616675 AV616675
37	210.4	6.6	750	11	BG585786 BG585786
38	208	6.5	851	11	BF969062 BF969062
39	206.4	6.5	803	11	BG298756 BG298756
40	204.4	6.4	787	11	BG584063 BG584063
41	199	6.3	854	13	A2682568 A2682568
42	197.8	6.2	402	11	H82550 H82550
43	197	6.2	541	10	A1722858 A1722858
44	196.4	6.2	2676	12	AK014319 AK014319
45	195.6	6.2	704	10	BE239652 BE239652

#### ALIGNMENTS

RESULT 1  
AL040762  
LOCUS DKF2P434C1815.r1 434 (synonym: htes3) Homo sapiens CDNA clone  
DEFINITION DKF2P434C1815 5', mRNA sequence.  
ACCESSION AL040762  
VERSION AL040762.1 GI:5409708  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 345)  
AUTHORS Bloeker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
EST (Bloeker, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Bloeker H

FEATURES  
source

MIPS  
Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the CDNA sequencing consortium of the  
German Genome Project.  
s1 sequence (DKF2P434C1815) is available at the RZPD in Berlin.  
This clone (DKF2P434C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY. Email: clone@rzd.de.

location/Qualifiers  
1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKF2P434C1815"

/clone\_1ib="434 (synonym: htes3)"  
/lissue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI"  
BASE COUNT 171 a 89 c 138 g 147 t  
ORIGIN

Query Match 17.2%: Score 545; DB 10; Length 545;  
Best Local Similarity 100.0%; Pred. NO. 2.7e-127;  
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 503 ccagatagaactcttcacacagctgatalataaactgatalataagaagaactgtg 562  
DB 1 CCAGTATAGTAACTTTCCACAGCTGATATAACTGATCAATGAGAGAACTGTG 60  
OY 563 aatttaaaatgttcttcaatlatcatcaagaaccatcatcaagatctgaaggtc 622  
DB 61 AATTTAAATGTTTCTTCAATTTATCCATCAAGACCATCTATCAAGATTCTGAAAGTTC 120  
OY 623 tgaattccagaattgaatctgagagagagctgccttggtcgttcaatgagagtgga 682  
DB 121 TGAATCTCAAAATTAAGTCTGAGAGACAGTCCGCTTGCTGCTCAATGCGAGTGGGA 180  
OY 683 agagtaagctgacacagctctcgaaggtatataatccgagatgagtgcttatacag 742  
DB 181 AGAGTACGATGACACACTCTCGAGAGGTTATATGATCCGAGATGATGCTTTATCATG 240  
OY 743 tgaatgagaatgacatcagaagctttaaagtgcgcatctcgagaccatctgagagtg 802  
DB 241 TGGATGAGATGACATCAGAGCTTTAATGTGCGCATTTACGAGACCATATTTGAGTGG 300  
OY 803 ttagtaagaagcctgtttcttcgagacacacatcaataaagtaagtaagagag 862  
DB 301 TTAACTCAAGACCTGTTTGTTCGGGACCAACATCATTAACAATATGAGTACAGAG 360  
OY 863 atgatgactgacatgagaagatgagagagcagaagaagaagaacatgcatgattta 922  
DB 361 ATGATGACATGATGAGAGATGAGAGAGACAGCAAGCAAGCAATGCGATATTTTA 420  
OY 923 tcatgagacttccataaatttaatacatctgtgaggggaaagaggtccaaatgtg 982  
DB 421 TCATGAGGATTTCTTAATTAATTAATGATGAGGGAAGAGAGCTCAAAATGAGTG 480  
OY 983 gaggagcaagaacagagatgcacatgctgcgttcaatgagaaacccaagattcga 1042  
DB 481 GAGGCGAAGAAACAGAGATGCGCAATGCTGCTGCTTACGAAACCCCAAGATTGCGA 540  
OY 1043 ttcta 1047  
DB 541 TTTTA 545  
RESULT 2  
AL520322 943 bp mRNA EST 13-FEB-2001  
LOCUS AL520322 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CSODB006Y15 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL520322 GI:12783815  
VERSION AL520322.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 943)  
AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry Cedex - France

Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source Location/Qualifiers  
1. .943  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CSODB006Y15"  
/clone\_1ib="LTI\_NFL004\_NBC2"  
/sex="male"  
/lissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 225 c 253 g 211 t 1 others  
ORIGIN

Query Match 12.2%: Score 386.6; DB 10; Length 943;  
Best Local Similarity 65.0%; Pred. NO. 4.2e-87;  
Matches 586; Conservative 0; Mismatches 314; Indels 1; Gaps 1;

OY 2270 taatttactgaatgcatgagagcagatgagcagcagaaagcgtcgttggctc 2329  
DB 6 TATTATTTCAGCTGTTGTTGGTCCATGGCCGTGGGCAAGTCAGTTCAATTGCTC 65  
OY 2230 ctgaatattccaaagccaatcggggctgcgcatctgttgcctgttgaaagaagac 2389  
DB 66 CTGACATATGCCAAAGCCAAATATCAACACCCACATCATATGATCAATTTGAAACCC 125  
OY 2390 caaatatagaagcggcagcgaagaaggaagaagcagaacatgtaaggaattag 2449  
DB 126 CTTGATTTGACAGCTTACAGCAGGAGGCTTAATGCGAACACATTGGAAGAAATGCA 185  
OY 2450 agttcgaagaagctcttcttcaatcagatgctgcagagatgttttccctccgtgct 2509  
DB 186 CATTTGGTGAAGTGTATTCACATCCACCCAGCCGAGCATCCAGTCTTACAGGAC 245  
OY 2510 tatccctcagatgagcagaagaagcagtagcattgtgggaggaagcggctgtgga 2569  
DB 246 TGAGCCTGAGGTAAGAAAGGCGCAGACGCTGCTGTGGGAGAGAGTGGTGGGA 305  
OY 2570 aaagcacttctgttcaacttctcagagactttagaccccggtcgaaggaagtcgt 2629  
DB 306 AGAGCACAGTGTCTCAGCTCTGAGGCGTTCTACGACCCCTTGGCAGGAAAGTGTGC 365  
OY 2630 ttgatgtgtgagatgcaaaagaattgaattgacagtggtcgttcccaaatgaaatcg 2689  
DB 366 TTGATGGCAAGAAATAAAGCGACTGATGTTACAGTGCTCCGAGCACACTTGCGGATCG 425  
OY 2690 ttccctcaagagcctgtgctcttcaactgagcacttgcctggaagaatgctatgtgga 2749  
DB 426 TGTCCAGAGAGCCCATCTCTGTTGACATGACCATTTGCTGGAAGAAATTTGAGAGCA 485  
OY 2750 acagccggtgtgtgctcattagatgagatcaagaagccgcaatgcaagcaatatcat 2809  
DB 486 ACAGCGGGGTGTGTCCAGAGAGAGATGCTGAGGGCAGCAAGAGGCGCAACTATCATG 545  
OY 2810 ctttattgaaggtctccctgagaataacaacacaaagtgtgactgaaggaagacagc 2869  
DB 546 CTTTCATGAGATGACATGCTGCTAATAATATACACATAAAGTGAAGACAAAGAACTAGC 605  
OY 2870 ttcttgccgccaagaacaagaactagctatgcaaggtccttctccaaacccaataa 2929  
DB 606 TCTCTGTGGCCAGAAACAGCATTTGCCATAGCTGTGCTGTTGTTGAGAGGCTCTATA 665  
OY 2930 ttattcttgatgagcacttcagccctcagataatgaaagatgagaaggtgtgtcagc 2989

```

Db      666  TTTTCCTTTGATGAAGCCAGCTGCTGTGATACAGAAAGTAAAGTTGTCACAG 725
Qy      2990  atgccccttataaagccaggaaggaagacatgcttagtgcctacacagctctctg 3049
Db      726  AAGCCCTGGACAAAGCCAGAGGAGCCGACCTGCACTTGTGTGATTTGCTACCCGCTTCCA 785
Qy      3050  caattcaagaacagatttgatagtg-gtctgcacaatggaagaataaggaacaaga 3108
Db      786  CCATCCAGAAATGACACTTAATATAGTGTGTGTTTCAGATGCGAGATCAAGAGCATGGC 845
Qy      3109  actcatcaagaagctctcggaagaatcgagacatataatttaagtttagtgaatgcacagta 3168
Db      846  ACGCATACACGACTGCTGTGCACAGAAAGCATCTATTTTCAATGTCAGTGTCCAGGCT 905
Qy      3169  g 3169
Db      906  g 906

```

## RESULT 3

```

AA243820 405 bp mRNA EST 06-AUG-1997
LOCUS z67906.r1 Soares_NHMPy_S1 Homo sapiens cDNA clone IMAGE:668506 5'
DEFINITION similar to SW:MDRL_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;
MRAA sequence.

```

```

ACCESSION AA243820
VERSION AA243820.1 GI:1874631
KEYWORDS EST.
SOURCE human.

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

## REFERENCE

```

1 (bases 1 to 405)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gaisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Maira,M., Martin,J., Moore,B.,
Schellenberg,K., Steploe,M., Tan,F., Theising,B., White,Y., Wyllie
,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)

```

## JOURNAL

```

COMMENT
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```

```

Email: estevan@wustl.edu

```

```

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1572 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham.

```

## FEATURES

## source

```

1..405
/organism="Homo sapiens"
/db_xref="GDB:5562473"
/db_xref="taxon:9606"
/clone="IMAGE:668506"
/clone_lib="Soares_NHMPy_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/notes="Organ: mixed (see below); Vector: PT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2BDHM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of 1 M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

```

## BASE COUNT

```

115 a 81 c 85 g 124 t
ORIGIN

```

```

Query Match 12.1%; Score 385.8; DB 10; Length 405;
Best Local Similarity 99.5%; Pred. No. 5.8e-87;
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      378  tcttgctgtcttccttagtgaatccatagcaattatgtcatgtgagcagccctca 437
Db      17  TCTCTGTCTTCTTTAGTGAATCATATGACACTTTTGCATTGGAGCAGCATCCCTCA 76
Qy      438  ctltgaaccttcgcaatagcccgagagctgccttccattatccaggltatgataa 497
Db      77  CTTGAACCTTCGCAATFAGCCGAGAGACTGCTTCATATTTCCAGGTATTGATFA 136
Qy      498  gaaccacatataagaacttttccacagctgatatataaccgtaatccatagaagaac 557
Db      137  GAACCCAGTATATGATTAACCTTTCCACACTGTGATTAACCTGAATCCATGAAGAAC 196
Qy      558  tctggaatttaaaatgcttcttcaattatccatcaagaaccatcatagaattctga 617
Db      197  TGTGGAATTTAAATATGTTCTTTCAATTATCCATGACACATCTATCAAGATTCTGAA 256
Qy      618  aggtctgaatctcagaatlaagctctgagagagacagctgccttgctgcctcaatgca 677
Db      257  AGGTCTGAATCTCAGAAATTAAGTCTGAGAGACAGTCGCTTGTCGTGCTCAATGCGAG 316
Qy      678  tgggaagagtaagtgtagtccagcttctcgagaggtatattatgacgagatagcttat 737
Db      317  TGGGAAGAGTACGCTAGTACGCTTCGACAGAGTATATGATCCGATGATGCTTTAT 376
Qy      738  catgtgtagtgaagaatgacatcagagctt 766
Db      377  CATGTGTGATGAGATGACATCAGACCTT 405

```

## RESULT 4

```

BF692596 559 bp mRNA EST 22-DEC-2000
LOCUS 602248949P1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
DEFINITION mRNA sequence.

```

```

ACCESSION BF692596
VERSION BF692596.1 GI:11978004
KEYWORDS EST.
SOURCE human.

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 559)
NIH-MGC http://mhc.nci.nih.gov/.

```

```

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

```

## JOURNAL

```

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCID/DTF

```

```

cDNA library Preparation: CLOUTERCH Laboratories, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

```

```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

```

```

Plate: L1CM1203 row: a column: 13
High quality sequence stop: 555.

```

```

Location/Qualifiers
1..559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4334100"
/clone_lib="NIH_MGC_62"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgccctggcc); Site_2: SfiI (ggccatcattggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'

```

## FEATURES

## source

BASE COUNT  
ORIGIN

BASE COUNT	263 a	203 c	242 g	185 t	1 others
ORIGIN					
Score 298.8:					
DR 11:					
length 894:					

```

419 ++++++ccccccctcctacgtttttaaacttcgaataaccgcaagaagctgccttcata 478
Query Match          9.4%; Score 298.8; DB 11; Length 894;
Best Local Similarity 66.6%; Pred. No. 7.5e-65;
Matches 471; Conservative 0; Mismatches 233; Indels 3; Gaps 3

```

[illegible][illegible]

QY 559 CAGAACACATCTACGAATCTGGAATTTAAGAATATTCACPTTACCATTCGCAA 217

218 AAGAGTTCAGATCTTGAAGGGCCCTCAACTGAGGTGAAGAGCCGGACAGACCGTGGCC 277  
 659 TGGTCGTCCTCAATGTCAGCTGGAGAGTACAGTATGTCAGCTCCGACAGGTATATG 718

**Dc** 719 atccgagatgatgccttattcatcggtgatgagaatgaacttcagagcttaaatgctgcgc 778  
**Dd** 278 TGGTTGGCAACAGTGGCTGTGGAAAAAGACAACCTGTCAGCGTGATGCAAAAGCCTTAGC 337

DB 338 ACCCCCTAAGATG6CATGCTGACGTATGACGACGACGACATCAGAACCATCATGTGTAGGT 397

Db 398 ATCTGAGGAGATCATTTGGTGTGGTAGTCAGGAACCTGTGCTGTTGGCACACCATG 457

Db	458	CCGAGAACCTTGCTATGGCCCGCAGAGATGTACACCATGATGATGATGATTTGAAGAAAGCTGTCA	517
Qy	899	gggaagcaatgcgatagtattatcatcggagtttcctaaataattatatacattggttag	958

Db	518	AGGAACCCATGCCCTTGA	CTTCATCATGAA	CTGCCCCACCA	TTTGAACACCC	CTGGTG	577
QY	959	gggaanaagagc	tcacatgagtcgagtcgagggcgagaacagagatcgcgaattgctcg	tgcc	1018		

Db 578 GTGAGAGAGGGGGCGACCTCAGTGGGGGGGACAGAAACGAGAAATCGC-ATTGGCCCCGGGGCCC 636

Qy 1019 tagttcgaaccccaagatctgatlltagatgaagctacgcttgcctcgatctgaa 1078

Db	637	TGATCCGCATCCCAAG-TGCCTTTCTGTGACGACGACACCTTAG-CTTGATATACAGAA	694
Qy	1079	gcaagtcagcgtgttcaagctgcactgcagagaagcgagcaaggtcgg	1125

Db 695 GTCTAAGCTGGGCTTCAGGCCGACCTAGATAGGGTACAAAGGCGCNG 741

RESOLUTION	0				
LOCUS	BG293345	726 bp	mRNA	EST	21-FEB-2001
DEFINITION	BG293345				
	602390738F1 NIH_MGC_94		MUS musculus cDNA clone IMAGE:402552		5'
	mRNA sequence.				

ACCESSION	BG293345
VERSION	BG293345.1
KEYWORDS	GI:13052943
SOURCE	EST.
ORGANISM	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov
TITLE	Tissue Procurement: The Cepko Laboratory
JOURNAL	cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LMNL0371 row: 1 column: 09 High quality sequence stop: 658.
FEATURES	Location/Qualifiers
source	1..726 /organism="Mus musculus" /db_xref="taxon:10090" /clone="IMAGE:4502552" /clone_1id="NIH_MGC_94" /tissue_type="retina" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: pCMV-SPORT6; Site:1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."
BASE COUNT	189 a     220 c     190 g     127 t
ORIGIN	
Query Match	8.7%; Score 275.2; DB 11; Length 726;
Best Local Similarity	68.3%; Pred. No. 7e-59;
Matches 396; Conservative 0; Mismatches 183; Indels 1; Gaps 1.	
QY 2540	tagcaattgtggygaagacagcgycitggyaaaagcaactctgttcaacttcgcagagac 2599
Dd	
3	TGGCCCTGTGTGGCGAGCATGTGGCTCGGGGAAGAGCACAAGTGTCACAGCTCGAAGCGT 62
QY 2600	tttatgaccgccgtgcaagagacaagtgcgtgttgtatgtgtgttgatgacaaagaattgaatg 2659
Dd	
63	TCTATGACCCTATGGCTGTGATCATGATGCTTTAGATGATCAAAAGAACATCAATG 122
QY 2660	tacagtgctccgttccccaaatagcaaatcgtttcccaccaagaagcctgtgctctcaacttga 2719
Dd	
123	TCCAGTGGCTCCGAGACTTAACTAGCGGCAATTGTGTCCCAAGAACCCATTCTTTGACTGCA 182
QY 2720	gcattcgttagaacaatcgcctatagtgtgaacaagccgtgtgtgtgcgaatagatatgatata 2779
Dd	
183	GCAATGCAGAGAACAATCCCTATGTGAGACAAACAGCCGGGTGTCGCTCATGTAGATG 242
QY 2780	aagaagccgcaaatgcaagaataatccattcttttatgtaaggtctccctgtgaatatata 2839
Dd	
243	TGAGGGCAGCCAAAGAGGCCAAACATCCACCCCCTTATCGAAGACGGTGGCCCCAAAATAATA 302
QY 2840	aacacaagcttggacgtgaagagacacagcttctgcgcgcagagaacaagaactagcta 2899
Dd	
303	ACACAAGAGTNGAGACACAAGGGGAGCGCAAGCTTCTGGGGGCCAAGAACAGAGATTTGCCA 362
QY 2900	ttgcaaggcctcttctccaanaaaacccaataatttatgtlttgatagagccacttcagccc 2959
Dd	
363	TGCGCCGAGCCCTCATTCAGACAGCTCGGGTCTACTGCTGTGATGAAGCACAGTCAGCTC 422
QY 2960	tcgataatgacacgtgaagaaggtgtgtcaagcatgtcccttgataaagcccaagcaggaaga 3019
Dd	

Dd	423	TGGATACTCGAGAGCTGAATAAGTTTGTCCAGGAAGCAACTGGACAACCGAGGAAAGGCCGA	482
Oy	3020	cattgcctagtgttcacccagcaggctctctgcaatccaagaacgaag-attgatagtgtt	3078
Dd	483	CCTGCATTGCCATCCTCACCCGCTCATCATCATCCAGAAGCGGACACTTGATCGGTG	542
Oy	3079	ctgccaatggaaataaagaacaagaactcatcaag	3118
Dd	543	ATTGGAACGGCAGGTCAAGAGCACCACCACCCACGAG	582

  

RESULT      7

BG248052					
LOCUS	BG248052	1019 bp	mRNA	EST	13-FEB-2001
DEFINITION	60235987P1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:448404				
ACCESSION	BG248052				
VERSION	BG248052.1 GI:12757867				
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1019) NIH-MGC http://mgs.nci.nih.gov// National Institutes of Health, Mammalian Gene Collection (MGC)				
AUTHORS	Unpublished (1999)				
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: rstraubs@mail.nih.gov				
COMMENT	Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. Genomic DNA Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLM10334 row: n column: 21 High quality sequence stop: 650.				

  

FEATURES

source	location/Qualifiers
	1..1019
	/organism="Mus musculus"
	/strain="FVB/N"
	/db_xref="taxon:10090"
	/clone="IMAGE:448404"
	/clone_lib="NCI CGAP Mam1"
	/tissue_type="tumor, biopsy sample"
	/dev_stage="3 months, virgin"
	/lab_host="DH10B"
	/note="Organ: mammary; Vector: PCMV-SPORE6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by life Technologies. Investigator providing samples: Gilbert Smith, NIH"

  

BASE COUNT

276 a	262 c	283 g	196 t	2 others
-------	-------	-------	-------	----------

  

ORIGIN

Query Match	8.6%; Score 273.2; DB 11; Length 1019;
Best Local Similarity	63.2%; Pred. No. 2.4e-58;
Matches 437; Conservative	0; Mismatches 253; Indels 2; Gaps 1

  

Oy	2437	gaaggaaatltagagtttcgagaagctctcttcattcacatgcccagaatgtttc	2496
Dd	4	GAAAGAAATGTGCAATTTTAGTGGACGCGTTCAACTAATCCACCGACCGACGATCCA	63
Oy	2497	atcccccggttatccctcaccatatttgaggcgaaagacagltgcatctgttggyaac	2556
Dd	64	GGCTTCAGGGGGCTCAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCTGTGGGCAC	123
Oy	2557	aagcgctgtgaaaagcactctgttcacctctgacagaagtatatgaccocgycaa	2616
Dd	124	AGTGCGTCGCGGAAGACACACTGGGTCCAGCTGCTCGAGCGCTTACAGACCCCATGGCT	183
Oy	2617	ggacaagtgtcgttgttatgtgtgtgtaatgaaaaaagaaatgaatglacagtgtccgttc	2676



D	b	184		GGATCAGCTGTTTTCAGATTGGCAAGAATAAAGCAACTGAATGTCCAGTGCGCTCCAGCA	243
OY		2677		caaatagcaatcgtctccataagaacgctgtgccttccaactcgacattgcttggaacatc	2736
D	b	244		CACGTGGGCATTGTGTGCCAAAGACCCTATTCTCTTTTACACGAGCATCGCAGAACATT	303
OY		2737		gccctatggtagcaaacagcccgttgtgtgcacctaaagtatatgaataaagaagccgaatgca	2796
D	b	304		GCTCACGGAGACAACACCGGGCTGTCTTTATGAGAGAAATTGTGAGGGCAGCCAMAGAG	363
OY		2797		gcaaatatccattcttttatgtaagytcctccctpgaaataacaacacaaagttagctg	2856
D	b	364		GCCAACATCCACCGATTGATGAGTACCTCCCTACCTGTAAATACAAACACCAAGTAGAGAC	423
OY		2857		aaaggagcacagccttctgtgcgcgcaaaaacaaagatactaattttaaggtctcttc	2916
D	b	424		AAAAGCACTCAGCTGGTGGGGGSCAAGACAGCGCATTCGCCATCGCACGCCGCTCTGTC	483
OY		2917		caaaaccoccaaatttattatgtgatagtgagccccactcaagccctcgaataagcaatgag	2976
D	b	484		AGACAGCCTCACCATTTTTACTTCTGTGAGAGAACACATCAGCTCTGGATTACAGAAAGTAA	543
OY		2977		aaggltggtltaaagcatgcgccttgataaagcagaagcyggaagagatgacctagtgtcact	3036
D	b	544		AAGTTGTCCAGGAAGGGCTGAGCAAAGAGCCAGGGGAAGCCGCACCTCATTTGTGTGCT	603
OY		3037		cacagggctctctgtaattcagaagacagatttgtagtggt--ctgcacaaatggaaga	3094
D	b	604		CACGCGTGTCCACCATCCCAAGACGGCGACTGTGCTGTGATTTCATAGAACGCGGAAG	663
OY		3095		taaaggaacaaagacatcataaagagctctgt 3126	
D	b	664		TCAAGGAGGACAGGGCACCACCAACAGCAAGTG 695	
R	E	S			
F	B	796582		944 bp mRNA EST 12-JAN-2001	
L	O			6022584463F1 NIH_MGC_85 Homo sapiens cDNA IMAGE:4341710 5'	
D	I			mRNA sequence.	
A	V			BF796582	
K	E			BF796582.1 GI:12101636	
W	S			EST.	
S				human.	
O				Homo sapiens	
R				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
P				Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
J				1 (bases 1 to 944)	
T				NIH-MGC http://mgc.ncl.nih.gov/	
N				National Institutes of Health, Mammalian Gene Collection (MGC)	
C				Unpublished (1999)	
O				Contact: Robert Strausberg, Ph.D.	
J				Email: rcapbs-re@mail.nih.gov	
C				Tissue Procurement: Louis Staudt, M.D., Ph.D.	
O				cDNA Library Preparation: Life Technologies, Inc.	
C				CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
D				DNA Sequencing by: Inocyte Genomics, Inc.	
I				Clone distribution: MGC clone distribution information can be	
F				found through the I.M.A.G.E. Consortium/LNL at:	
E				http://image.llnl.gov	
E				Plate: LLM955 row: n column: 15	
E				High quality sequence start: 9	
E				High quality sequence stop: 669.	
E				Location/Qualifiers	
E				1..944	
F					
E					

Site 2: Salty. Cloned unidirectionally; oligo-dT primed  
Average insert size 1.867 kb. Library enriched for  
full-length clones and constructed by Life Technologies  
Note: this is a NIH-MGC Library."

Query Match	8.5%	Score 269	DB 11	Length 944
Best Local Similarity	65.4%	Pred. No. 2,7e-57		
Matches 409	Conservative	0	Mismatches 215	Indels 1
				Gaps 1
Qy 2335	tattccaaagccaatctcggggctctgcgcatcgtttgcttctgttggaaaagaaaccaat	2394		
Db 39	TATGCTAAAGCTAAGCTGTCTGCAAGCCACTTATTCATGCTGTTTGAAGAACAACCTCTG	98		
Qy 2395	atagacagccgacgtcaagaagggaaaaagccagacacacatgtgaagggatlttagtltt	2454		
Db 99	ATTGACACGCTACAGTGAAGAGGGGCTGAAAGCCGTGTAATTAATTTGAAGAAATAATACATTT	158		
Qy 2455	cgagaagctcttcttcttattccatctgtccagatgtttccalcctccggtctaac	2514		
Db 159	AATGAAAGCTGTGTTTCAACTATTCACCCGAGCAAAAGTGCAGTGTCTCAGGGCTGAGC	218		
Qy 2515	ctcagatctgagcgagaagaacacagtagcatctgttgggagcaagcgtctgtggaaaaagc	2574		
Db 219	CTGGAAGTGAAGAAAGGCCAGACACTACCTCTGTGGGCGACAGTGCCTGTGGGAAGAGC	278		
Qy 2575	actctgttcaactctctgcagagactltgaaccccggtgcaaggaagaaggtctgttcat	2634		
Db 279	ACGGTGTCTCAGCTCTCTGAGCGGCTTACGACCCCTTGGCGGGACAGTGTCTTCGAT	338		
Qy 2655	ggttctgtatgtcaaaaagaattgaatgttaagtgatggtctccgttccaaatgacaatctctc	2694		
Db 339	GGTCAAGAAGGAAGAAATCTCAATGTCCAGTGGCTCAAGACTCAACTGTGGATCTGTCT	398		
Qy 2695	caagagccctgtgctcttcaactctgcagacatctgtgagaacatcgccatctgtgtgacaacagc	2754		
Db 399	CAGAGGCTTATTCATTTGACTGGAGCATTTGCCAGAAATTTGGCTATGGAACAAACAGC	458		
Qy 2755	cgttctgtgccattagatgtagatcaagaagccgcaaatgtcaagcaaatatcatctcttt	2814		
Db 459	CGGTTGTATTCACAGGAAGAAATCTGTGATCGAGCCAAAGGTGCCAATCATCTCTTC	518		
Qy 2815	attgaagcttccctcgagaaatacaacaacagaagtgtgacgtgaagaagagcagacttct	2874		
Db 519	ATCGAGACGTTACCCCAAAATATGTAAACAGAGTGGAGATTAAGGGACACTCAGCTTCA	578		
Qy 2875	ggcgcgcgaagaacaagaactagctatltgcagaggtctcttccaaaaaaccc-aaaatltt	2933		
Db 579	GGAGGTACAAACAGAGGATTTGCTATTGTCCCGAGCCCTCATCAGACAAACCTCAAAATCT	638		
Qy 2934	attgttgaatgagccacttcaagcc	2958		
Db 639	CCTGTTGGATGAAGCTACATCAGCC	663		
RESULT 9				
LOCUS	AVT09991	785 bp	mRNA	EST
DEFINITION	AVT09991 Cu Homo sapiens cDNA clone CUAJJA09 5', mRNA sequence.			09-OCT-2000
ACCESSION	AVT09991			
VERSION	AVT09991.1	GI:10728272		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:43417710"
/tissue_type="lymphoma, cell_line"
/lab_host="DH10B (phage-resistant)"
/name="Organ: lymph; Vector: pCMV-Sport6; Site:1; NotI;

```

TITLE Homo sapiens CDNA Cu clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1. .785  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CuADJ09"  
/clone\_id="Cu"  
/tissue\_type="adrenal cortico adenoma for Cushing's  
syndrome"  
/dev\_stage="Adult"  
/note="Vector: pBluescript sk(-)"

BASE COUNT 210 a 180 c 211 g 183 t 1 others  
ORIGIN

Query Match 8.3%; Score 262.2; DB 10; Length 785;  
Best Local Similarity 64.0%; Pred. No. 1.4e-55;  
Matches 458; Conservative 0; Mismatches 254; Indels 4; Gaps 4;

OY 2260 ggcagtcacatagatgacagcgcagtcacagaggaagaaagcagacacatgtgaa 2439  
DB 133 GAAAAAACCCCTTGTATGACAGCTACAGCACGGAAGGCTTATGCGCAACACATTGGAA 252  
OY 2440 gggaaattagaagtttcggaagtccttcttcttcaatgctgcgcacagatgtttcacc 2499  
DB 253 GGAATGTGACATTTGGGGAAGTGTATTTCACTATCCACCCGACGACATCCCAAG 312  
OY 2500 ctccgtgacctatccctcaagatgagcaggaagaaagacagtagcatgttgaggagcagc 2559  
DB 313 CTTTACGGGACTGAGCTGAGGTGAAGAGGCGCCAGACGCTGCTGTGGGCAAGCACT 372  
OY 2560 ggcgtgggaaagacattcgttcaactctgcagagactttagaccgcgtgcaagga 2619  
DB 373 GCGTGTGGAGAGCAGACAGTGTCCAGCTCCGAGCGGTTTACGACCCCTTGGCAGGG 432  
OY 2620 caagagctgttgaagtggtgagatgcaaaagatgaaatgacgtggtccgtccca 2679  
DB 433 AAGAGTGCTGTGATGAGCAAGAAATTAAGCAGCTGAATGTTCAAGTGGCTCCGACAGC 492  
OY 2680 atagcaatcgttctcacaagagcctgtgtcctcaactgcagcagcattgtagaagatgcgc 2739  
DB 493 CTGGGCACTGCTGTCCAGAGGCCATCTGTTGACTGACGAGCATTTGAGAAAGATTGGCC 552  
OY 2740 tatgtgacaagagc-gtgtgtgcatctagatagatcaagaagccgcaaaagcagc 2798  
DB 553 TATGGAGACACAGCCGGGGTGTGTACAGGAAGATTTGAGGGGCGCAAAAGAGAGC 612  
OY 2799 aaatcattcttattgaaggtctccctgagaataacaacacagaagtgtgagatga 2858  
DB 613 CAACATACATGCTTCATGAGACTCACTGCT-ATAAATATACACTAAACTNAGAGACA 671  
OY 2859 agagacacagcttctgagcgcagaaacaaagactagctattgcaaggctcttccca 2918  
DB 672 AGGACCTAGCTCTGTGGGGCGCAGAAACACACA-TTGGCATATGCTTGGCCCTGGTAA 730  
OY 2919 aaaccacaatattatgttgatgagggccactcagccctcgataagaagtg 2974

DB 721 CAGACTTATAA-TTTGTTTTGGATGAGCCACGCTGAGTGTGATACAGAAATG 785

RESULT 10  
LOCUS BC587938 780 bp mRNA EST 11-APR-2001  
DEFINITION EST489713 MHAM Medicago truncatula/Glomsus versiforme mixed EST  
library cDNA clone pMHAM-51L19 5' end, mRNA sequence.  
ACCESSION BC587938  
VERSION BC587938.1 GI:13603002  
KEYWORDS  
SOURCE Medicago truncatula/Glomsus versiforme mixed EST.  
ORGANISM Medicago truncatula/Glomsus versiforme mixed EST library.  
REFERENCE Eukaryota, mixed EST libraries.  
AUTHORS Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Uteerback,T., Cho,J.  
and Fraser,C.M.  
ESTs from roots of Medicago truncatula after colonization with  
Glomsus versiforme, 2001  
Unpublished (2001)  
CONTACT: Harrison M.J.  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73401  
Tel: 580-223-5810  
Fax: 580-221-7380  
Email: mjharrison@noble.org  
Noble EST name: N38701le TIGR sequence name: MTDNM70TK More  
information is available at: <http://www.medicago.org>  
Seq primer: SKmod (CTA GAA CTA gtc gat CC).

FEATURES  
source  
1. .780  
/organism="Medicago truncatula/Glomsus versiforme mixed EST  
library"  
/cultivar="Medicago truncatula genotype A17"  
/db\_xref="taxon:119092"  
/clone="pMHAM-51L19"  
/clone\_id="MHAM"  
/tissue\_type="Roots colonized with Glomsus versiforme"  
/dev\_stage="Roots harvested at 10, 17, 22, 31 and 38 days  
post-inoculation with Glomsus versiforme. The library was  
made from a mixture of RNA from each of these stages."  
/lab\_host="E. coli strain XL0R"  
/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:  
XhoI; cDNA was prepared from polyA+ enriched RNA from  
roots harvested at 10, 17, 22, 31 and 38 days  
post-inoculation with Glomsus versiforme. The cDNA was  
directionally ligated into the UniZap XR vector from  
Stratagene and packaged using Gigapack III Cold packaging  
extracts. Plasmids containing cDNA inserts were excised  
from the recombinant lambda-Zap phage using Ex-assist  
helper phage and propagated in XL0R cells."

BASE COUNT 234 a 156 c 184 g 206 t  
ORIGIN

Query Match 8.0%; Score 253.2; DB 11; Length 780;  
Best Local Similarity 58.5%; Pred. No. 2.7e-53;  
Matches 441; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

OY 419 ttggagcagcagctccctcacttgaaccttcgcaatagccgagagagctgccttcata 478  
DB 4 TGGGGCAATCTGCACCGACATGCTGCAATTAACAAAGCTGAGTGGGCTGTAA 63  
OY 479 ttctcaggtatgataaagaccagatagataactttcccaagctgataaac 538  
DB 64 TTTTCGGATATTGATTCACCGCCGTGATATGATGAACAGTGAATTTGGATTGAA 123  
OY 539 ctgaatccatagaaggaactgtggaatttaaaatgtttcttcaattatccatcaagc 598  
DB 124 TAGAGACACTTACTGAGACTTGTGAACCTGAATAATGTGACTTCTTATTCATCAAGAC 183

OY		599	cacatcaagaatttcgtgaagagctcgaatcccaaatgaatctcggaagagcagtcgacct	658
Db		184	CTGAAGTTCCTCAATCTCCTCAATGATTTCCTCGTTGAGTGTTCCGCCGGAAAGCATACGTT	243
OY		659	tgcgtgcgtccaatggcagtcgggaagtagtacgtagtaccgctcttcgcagagtgataag	718
Db		244	TAGTGTGATGACGAGCGGCTCTGGCAAGACACTGTGTCTTCTTAATGAGAGATTCTAAT	303
OY		719	atccgagataggcgtttalcalcgvgagatgaatgatcatacaagaagctttaagtgcgc	778
Db		304	ATCCACTTCACGAGACAATAATGTTGGATGGGCATGATGATTAABCTTTGAAGCTTAAT	363
OY		779	attalcgagaccatatctggatgggtiaagtaagaagcctgtttctgctcggaaccaacaa	838
Db		384	GATTGAGGCAACAATAATAGACTAGTAGACCACCAAGAACCCTGCTTTGTTGCCACACGATTC	423
OY		839	gtacaataatacaagatagtagcagagatgtagtgcactgatgaagagatggaagagcaagaa	898
Db		424	GAGAAAATATACTCTGTGGGAGGCGCTGATGCAAAACAGAGTTGAGATTGAAGAAGCTTGCTA	483
OY		899	gggaagcacaatgcgtatgatattatcatcagagagttcccataaattatacatgtgtag	958
Db		484	GSGTTGCTAAATGSCGTATTCATTCATCACATCAACAACCTTCGGAAGGCTTTGAACCTAGTAG	543
OY		959	gggaagaagsgctcaaatgagtcgagggcgagaacacagagatgcgaatctgcgtgcct	1018
Db		544	GAGAAAGAGGACTACACACTTCTGTGAGGAGCAAAAAACAAGATATGACATATACCAAGGCGAA	603
OY		1019	tagttcgaagacccccagaagtttgtatatttagtttgaggtgcagtcgtgcccttgatcaagaa	1078
Db		604	TGCTAAAAAATTCACCAATTTCTCTCTCGATGAGAGCAACAAGTCATTTGAGCTCTTGAT	663
OY		1079	gcaagtcagctgtltcaagctgcactcgtgagaagcgcgagaagaagtgtagcctaacaatcgcyg	1138
Db		664	CAGAAACACTGCTGGTGCAGAAAGACTTGTGACCGGTTTCATGATTGGCAGACAACTCTGTAA	723
OY		1139	tagcacaccgagctttctactatctgaagtgcaga	1172
Db		724	TTGCTCATCGCCTCTCCACAAATTCGCAAAAGCTGA	757
RESULT		11		
LOCUS		BE749379	560 bp	mRNA EST 25-APR-2001
DEFINITION		200194 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.		
ACCESSION		BE749379		
VERSION		BE749379.1	GI:10163371	
KEYWORDS		EST.		
SOURCE		cow.		
ORGANISM		Bos taurus		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
		Bovidae; Bovinae; Bos.		
		1 (bases 1 to 560)		
		Smith,T.P.L., Grosses,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,		
		Casas,E., Wray,J.E., White,J., Cho,J., Fahnenrugg,S.C., Bennett,		
		G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-mckown,C.G.,		
		Petee,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,U. and		
		Keele,J.W.		
TITLE		Sequence evaluation of four pooled-tissue normalized bovine cDNA		
JOURNAL		libraries and construction of a gene index for cattle		
MEDLINE		Genome Res. 11 (4), 626-630 (2001)		
COMMENT		21180013		
		Contact: Smith TPL		
		USDA, ARS, US Meat Animal Research Center		
		PO Box 166, Clay Center, NE 68933-0166, USA		
		Tel.: 402 762 4366		
		Fax: 402 762 4390		
		Email: smiththe@mail.marc.usda.gov		
		Single pass sequencing. Bases called and alt_trimmed with phred		
		v0.980904.e. Vector identified by cross_match with the -mismatch		
		and -minmatch 12 options.		
		PCR primers		

FEATURES		source		/organism="Bos taurus"		/db_xref="taxon:9913"		/clone_lib="MARC 480V"		/tissue_type="pooled"		/lab_host="DH10B"		/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from day 20 and day 40 embryos."	
Location/Qualifiers		1..560													
BASE COUNT	154 a	153 c	146 g	107 t											
ORIGIN															
Query Match	7.9%; Score 252.4; DB 10; Length 560;														
Best Local Similarity	67.5%; Pred. No. 4e-53;														
Matches	355; Conservative	0;	Mismatches	171;	Indels	0;	Gaps	0;							
Oy	2646	aaagaattgaatgtaacagagtgctcgttcgtcccaaatagcaatcgttcctccaagagcgtg	2705												
Db	1	AAAGAGCTCAACGTCACGTCAGTGGCTCAGAGCCCAACTTGGAAATCGTGTTCAGAGAGCCCT	60												
Oy	2706	gctctcaactcagcatctgcgtagaacaatcgcctatggtgtagaacaacgcgtgtgtgc	2765												
Db	61	CCTGTTGACTGCACACATTGGCCGACAACTTCCCTATGGGGAACAAGCCGGGCGTGTAC	120												
Oy	2766	attagatgagatcaaaagaagccgcgaatgcagcaaatatccattctttattgaagctc	2825												
Db	121	CATGCTGAAATTTGAGAGCGCAGCCAAAGCACCAACATCCATCTTATTGAGACCTT	180												
Oy	2826	ccctgagaatacaacacacaagaattggactgaagaagagacagcttctgcgcgcaga	2885												
Db	161	GCCCCACAAATTTGAAACAGAGTGGAGATTAAGGGGACTCAGCTTCCGGGGGACAGA	240												
Oy	2886	acaaagactagctattgtcaaggctctctccaaaaaocccaaatttatgtgtgata	2945												
Db	241	ACAGAGGATTTGCTATTGGCCCGAGCCCTCATCCGACACCCCGCATCCTACTCTGGATGA	300												
Oy	2946	ggccactcagccctcgatgaatgacagtagaagaattgtgtcagatgccttgataaagc	3005												
Db	301	AGCCACGTCAGCTGACTGATCTGAAGTAGAAGATTGTCCAAAGAGCCCTGGACAAAGC	360												
Oy	3006	caggaacgggaagaatgcctagtggtcactcacagagctctctgcgaattcaagaacgaga	3065												
Db	361	CCGAGAGAAGGCCACCTCGATCGTATCCGTCACCGGCTGTCCACCATCTCAAAACGACA	420												
Oy	3066	tttgatagtgtcttcgcacaatggtgaagaatataagaagaacgaactcatcaagagctct	3125												
Db	421	CTTGTATAGTGGTATGGAATAATGGCAGAGTCAGGGAGCAACGACACACACAGCAGCTGT	480												
Oy	3126	gagaatcgcagcatatatttaagttagttagatgcacagctagtg	3171												
Db	481	GGCACAGAAGGATCTATTTCACCATGTCTCACTGTCCAGGCTGGG	526												
RESULT 12															
LOCUS	BH139685	886 bp	DNA	GSS	07-AUG-2001										
DEFINITION	ENTNNA47TR Entamoeba histolytica sheared DNA Entamoeba histolytica genomic, DNA sequence.														
ACCESSION	BH139685														
VERSION	BH139685.1	GI:15098746													
KEYWORDS	GSS.														
SOURCE	Entamoeba histolytica.														
ORGANISM	Eukaryota; Entamoebidae; Entamoeba.														
REFERENCE	1 (bases 1 to 886)														
AUTHORS	Loftus,B., Wang,Z., Van Aken,S. and Frazer,C.														
TITLE	Determination of clone end sequences from Entamoeba histolytica														

JOURNAL  
COMMENT

HMI:IMSS sheared DNA library (2001)  
Unpublished (2001)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: entaefr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: Shotgun  
High quality sequence start: 18  
High quality sequence stop: 851.  
Location/Qualifiers

## FEATURES

SOURCE

1. 886  
/organism="Entamoeba histolytica"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica sheared DNA"  
/note="Vector: pHOsi; Site\_1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
light size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaubin and B.  
Bartell, Oxford University Press, 1999)."

BASE COUNT 348 a 105 c 176 g 257 t  
ORIGIN

Query Match 7.9%; Score 252; DB 13; Length 886;  
Best Local Similarity 58.5%; Pred. No. 5.5e-53;  
Matches 480; Conservative 0; Mismatches 330; Indels 10; Gaps 2;

Oy 461 gagagagcgccttcatttcacaggtattgataaagaaacccagatagtaacttt 520  
Db 12 GAAGAAGAGCAAGAAAGATCTATGATGATTTGAT-AGAAAACCACTATGATGTTATA 70  
Oy 521 ccacagctgatalaataacatcatcatalagaagaacgttgaaatttaaaatgttctt 580  
Db 71 GTGAAGAGAGTGAAGAAACATTTATGATGTTAAAGTGCAATTTGAATTTAAAGCATTTGTT 130  
Oy 581 tcaattatcatcaagaacatcatcatalagaagaacgttgaaatttgaatcagaatgaat 640  
Db 131 TCAGATATCCCAACAGACAGACATTTCTGTTGAAAGGATTTTCATTTCAAGATGAGAAC 190  
Oy 641 ctgagagagacgtgccttgctgcgttcacatgagcagtgagagagatcgcgtaagtcagc 700  
Db 191 AAGGAAAACGTTGTCATTTAGTAGAGCATCAGATGTGTAATCAACATCAGTTTCACT 250  
Oy 701 ttctgcagaggttlatatgacccgagatgagtcctatcatcagtgatgagatgacatca 760  
Db 251 TGATGAAAGATTTTATGATGCAACATGAGATGATTTATTTAGATGACATTAATATCA 310  
Oy 761 gagctttaaagtcgagcattatcgagacatattgagtgatgagtaagaagcctgttt 820  
Db 311 AAGATTTAAATATTCATTTCTTAAGACATGCAATTTGGAATGCTAGACAGAACCATAT 370  
Oy 821 tcttcgagagacccatcagtaatacaatcaatgagtgatgagtgagtgagtgagtgag 871  
Db 371 TATTTGCTGAAGAGTGTATTTAGATTAATTTAGAGAGAGTACCTTAAGAGTTGAAGTAA 430  
Oy 872 ctgagtaagagatgagagagagcaagaaggaagaacatgagtgatattatcatcagaggt 931  
Db 431 GTAATGAACAATTTATGCTGCTGCTAAATGCAATGCAATGACATGACCTTTATTTCAGCAA 490

Oy 932 ttccataaatttaatacattgtaggggaaagagagcctcaatgagtgagggcaga 991  
Db 491 TGCCAGAGGATATATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550  
Oy 992 aacagagagatgcgaattgctgcgttcagtgatgagtaaacaccccaagatcttgatgag 1051  
Db 551 AACAAAGAAATGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610  
Oy 1052 aggtcagtcgtcccttgatcagagagagagagagagagagagagagagagagagag 1111  
Db 611 AAGCTACATCAGCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670  
Oy 1112 cgaag 1171  
Db 671 CAGCAAG 730  
Oy 1172 attgattgtagccctaaagagatgagatgagatgagatgagatgagatgagatgagat 1231  
Db 731 ATCAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790  
Oy 1232 tggcaaacgaggtctatattatcattgagtgagtgagtgagtgagtgagtgagtgag 1271  
Db 791 TAGATTTGAAGGATTTAT 830

## RESULT 13

BF969667 981 bp mRNA EST 22-JAN-2001

LOCUS 602272046P1 NIH\_MGC\_84 Homo sapiens cDNA clone IMAGE:4360090 5',

DEFINITION mRNA sequence.

ACCESSION BF969667 .GI:12336882

VERSION BF969667.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 981)

AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/MLN at:

http://image.llnl.gov

Plate: LLM10000 row: 1 column: 11

High quality sequence stop: 670.

Location/Qualifiers

1. 981

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="IMAGE:4360090"

/clone\_lib="NIH\_MGC\_84"

/tisssue\_type="adrenal cortex carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: adrenal gland; Vector: pCMV-SPORE6; Site\_1:

NotI; Site\_2: SalI; Cloned unidirectionally; Oligo-dT

primed. Average insert size 1.229 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_MGC Library."

BASE COUNT 267 a 194 c 250 g 270 t

ORIGIN

Query Match 7.9%; Score 249.4; DB 11; Length 981;

Best Local Similarity 61.4%; Pred. No. 2.5e-52;

Matches 400; Conservative 0; Mismatches 251; Indels 0; Gaps 0;



```

QY 1086 agctgtcaagctgacgtgagaaagcgagcaaggtcgcgatacatcgtgtgagaca 1145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 188 GATTGTACAGACAGCACTTGAAAAAGCTAGTCAAGAGAACAAACATTTGTTGACACA 129
QY 1146 ccgacttctactacttcgaagtcgagattgatt 1179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 128 TAGATTAACTGTTAGAAATGCAAGTAGAATT 95

RESULT 15
BF313560 795 bp mRNA EST 21-NOV-2000
LOCUS 601900192F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4129222 5',
DEFINITION mRNA sequence.
ACCESSION BF313560
VERSION BF313560.1 GI:11261583
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 795)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: LDCM1023 row: h column: 23
High quality sequence stop: 674.
FEATURES
source 1..795
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4129222"
    /clone_1lb="NIH_MGC_19"
    /tissue_type="neuroblastoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
    EcoRI; cDNA made by oligo-dt priming. Directionally
    cloned into EcoRI/XhoI sites using the following 5'
    adaptor: GGCACGAG(G). Library constructed by Ling Hong
    in the laboratory of Gerald M. Rubin (University of
    California, Berkeley) using ZAP-CDNA synthesis kit
    (Stratagene) and Superscript II RT (Life Technologies).
    Note: this is a NIH-MGC Library."
BASE COUNT 213 a 189 c 236 g 157 t
ORIGIN

Query Match 7.8%; Score 249; DB 11; Length 795;
Best Local Similarity 62.1%; Pred. No. 3.1e-52;
Matches 479; Conservative 0; Mismatches 280; Indels 12; Gaps 5;

```

```

DB 178 TGGAGAGACACAGTGTGTCACGCTCTTGAGCGGTTTACGACGCCCTTGCGAGGAAAGT 237
QY 2625 gctgtttgattggtgtgtgagatgcaagaatgtatgtacagtgctccgttcccaatgac 2684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 238 GCTGCTGTATGGCAAAAGAAATTAAGCGACTGAATGTTCAGTGGTCGACGACACCTGGG 297
QY 2685 aatcgttcccaagaagccgtgtgcttcccaatgcaagatgctgtgtggaacatgctatgg 2744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 298 CATCGTGTCCAGGAGGCCCATCTCTGTTGACTGCAACATTCGTGAGAACATTGCTATGG 357
QY 2745 tgacaacagccgtgtgtgtgccaatagatgaga tcaagaagccgcgaatgcaagaatat 2804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 358 AGACAACAGCCGCGGTGTGTACAGAGAGATCGTAGAGCGCAGCAAGAGGACCAACAT 417
QY 2805 ccattccttattgaaagtgctccctgagaataacacacaagaattgactgaaagagac 2864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 ACATGCTTCATCGAGTCACTGCTCTTAATTAATATGACACTAAAGTAGGAGACAAAGGAAAC 477
QY 2865 acaagcttctgagcgccgcaagaacaagaactaactgttgcaagggctcttcccaaaac 2924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 TCAGCTCTCTGTGGCCGCAAAACGCAATGCCATAGCTCGTGCCCTTGTTAGACAGCC 537
QY 2925 caaatattattgtgtgagatgagccacttcaagccctcgataatgacagtgaaagtg-9 2983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 TCATATTTCCT--TTGATGAAGCCACGTCACCTCTGTGATACGAAGTGAAGAAAGTGCG 595
QY 2984 ttcaagatgcccctgtataaagccgaagacggaagacatgacctagtggtcactcaaggc 3043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 536 TCCAAGAAAGCCCTGGACAAAGCCA---GAGAAAGCGACCTCATGTGTGTACCGCC 651
QY 3044 tctcgcatttgaagaagcagaatttgatagtggtctgcacatggaataaagaataaac 3103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 652 TGTCCACCACTCCAGATGAGACTT--ATATGTGTGTTTCAGATGCGACAGTCCAGGGGG 709
QY 3104 aaggaactcatcaagagctcctctgagaatacgacatatattttaagtag 3154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 710 CATGACGCGATCAGAGGTGCTGACAGAAAGGCTCATTTTTCACGGTACG 760

```

Search completed: April 22, 2002, 18:32:03  
Job time: 12216 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 23:02:24 ; Search time 886.72 Seconds  
(without alignments)  
3071.682 Million cell updates/sec

Title: US-09-873-409-12

Perfect score: 3177  
Sequence: 1 atggtcattcattgaccag.....atgcacagtcagtcagtgta 3177

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq.1101:\*

- 1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1130	35.6	AAD03488	Dog P-glycoprotein
2	1130	35.6	AAD03504	Dog P-glycoprotein
3	1130	35.6	AAD03505	Dog P-glycoprotein
4	1130	35.6	AAD03506	Dog P-glycoprotein
5	1126.8	35.5	AAD03489	Dog P-glycoprotein
6	1125.8	35.4	AAD052726	Sequence of human
7	1124	35.4	AAH57442	Human intestine ce
8	1124	35.4	AAQ72872	Human multidrug re
9	1124	35.4	AAZ94738	Human ATP binding
10	1124	35.4	AAW70752	Sequence of human
11	1124	35.4	AAW13394	Hybrid vector pSP-

12	1122.4	35.3	3860	21	AAZ49332	Human wild-type mu
13	1122.4	35.3	3860	21	AAZ49333	Human G185V mutant
14	1122.4	35.3	4264	19	AAV66533	Mutated human P-gl
15	1122.4	35.3	4264	19	AAV66534	Mutated human P-gl
16	1122.4	35.3	8630	21	AAZ24042	Retroviral vector
17	1120.8	35.3	8630	21	AAZ24041	Retroviral M4 mdr-
18	1119.2	35.2	4378	11	AAQ04522	Multidrug Resistan
19	1114.4	35.1	4669	19	AAV32645	Human P-glycoprote
20	1112.8	35.0	4186	22	AAV86127	Cynomolgous monke
21	1112.8	35.0	4195	22	AAV86128	Cynomolgous monke
22	1108	34.9	3988	21	AAZ88973	Human MDR-1 DNA.
23	1097	34.5	3924	21	AAZ94742	Human ATP binding
24	1097	34.5	3924	21	AAZ88974	Human MDR-3 DNA.
25	1045.6	32.9	4369	21	AAZ52047	Rat multidrug resi
26	1045.6	32.9	4425	21	AAZ52048	Rat multidrug resi
27	1043.4	32.8	4189	21	AAZ49334	Murine multidrug r
28	1043.4	32.8	4313	14	AAQ38950	Mouse multidrug re
29	1040.6	32.8	4233	21	AAZ90198	Rat mdr1b2 multits
30	1040.6	32.8	4233	22	AAZ27498	Murine multidrug r
31	1028	32.4	4788	21	AAZ49335	Multidrug-resistan
32	856	26.9	2726	15	AAQ70807	Multidrug-resistan
33	856	26.9	2726	15	AAQ70816	Multidrug-resistan
34	856	26.9	2726	18	AAZ43322	Multidrug resistan
35	721.4	22.7	4776	21	AAZ94744	Human ATP binding
36	506	15.9	4175	20	AAV69392	H. contortus RGP-A
37	436.4	13.7	4051	21	AAQ09019	A. thaliana ATPAC
38	433.4	13.6	4102	21	AAZ45942	Arabidopsis thalia
39	426.6	13.4	4002	21	AAZ85823	CDNA encoding mult
40	386.6	12.2	4047	20	AAZ60201	CDNA encoding the
41	383.8	12.1	2698	20	AAV69394	H. contortus RGP-A
42	349.6	11.0	3512	19	AAV69395	H. contortus RGP-O
43	331.4	10.4	4224	19	AAV42347	CDNA encoding a mu
44	322.6	10.2	8777	22	AAV83399	P. chrysogenum ABC
45	322.6	10.2	8777	22	AAV83400	P. chrysogenum ABC

#### ALIGNMENTS

RESULT 1	
ID	AAD03488
XX	AAD03488 standard; cDNA; 4279 BP.
XX	AAD03488;
XX	13-JUN-2001 (first entry)
XX	Dog P-glycoprotein (PGP) cDNA #1.
XX	Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX	drug bioavailability; transgenic animal; genetic model; ss.
XX	Canis familiaris.
XX	
XX	Key
XX	Location/Qualifiers
XX	17..3862
XX	/*tag= a
XX	/product= "Dog P-glycoprotein (PGP) #1"
XX	
XX	WO200123540-A2.
XX	
XX	05-APR-2001.
XX	
XX	28-SEP-2000; 2000MO-US26767.
XX	
XX	28-SEP-1999; 99US-0156510.
XX	
XX	(GENT-) GENTEST CORP.
XX	
XX	Stocker PJ, Steinel-crespi DT, Crespi CL, Relf TC, Patten CJ;
XX	WPI: 2001-235373/24.
XX	P-PSDB; AAEO0303.



XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT PGP inhibitors -  
XX  
XX  
PS Claim 3: Page 58-63; 11pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity, by  
CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This  
CC sequence is also referred as Genotype C cDNA. The  
CC PGP enzyme functions as an efflux pump exporting small molecules  
CC across the cell membrane. This enzyme is a member of the ABC  
CC transporter family.  
XX  
XX

Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Query Match 35.6%; Score 1130; DB 22; Length 4279;  
Best Local Similarity 62.4%; Pred. No. 1,4e-295;  
Matches 1879; Conservative 0; Mismatches 1085; Indels 48; Gaps 5;

OY 200 ttgttaagttatcaacgaatctcaagaatgcaagatttggtacataaagaagctatag 259  
DB 843 ttcaaaagttcaacaaataattagaagaagctaaagaattggtatagaagaagctatca 902  
OY 260 ctcaaaagttctctgtgctgt 319  
DB 903 cggcaacattctctatgt 962  
OY 320 ttgtgttgtaacctctctgt 379  
DB 963 tctgtgttgtaacctctctgt 1016  
OY 380 ttgtgttgtaacctctctgt 439  
DB 1017 tcaatgtctctctctgt 1076  
OY 440 ttgaaccttcgaatagaccggaagagctgcttcatatttccaggtatgataaga 499  
DB 1077 ttgaagattttgcaaacggaagagagcttgaatcttcaagataatgataaga 1136  
OY 500 aaccagatatagaatactttccacagctggtatataaactggaatcattagaagaatg 559  
DB 1137 aaccagatatagaatcttgaagagagcttgaatataaactggaatataaagaat 1196  
OY 560 ttgaattataaagttcttcttcaattatcatcaagaaccctcatagaagtttgaag 619  
DB 1197 ttgaattataaagttcttcttcaattatcatcaagaaccctcatagaagtttgaag 1256  
OY 620 gtccgaattcgaatataagttctggaagagagctgctgtgtgtgtgtgtgtgtgtgtgt 679  
DB 1257 gtcccaacctggaagttcagaggtgtggaagagagctggtgtgtgtgtgtgtgtgtgtgt 1316  
OY 680 ggaagagtagagttgctcagcttctgcaagaggtatataagtcggaatgagttatca 739  
DB 1317 ggaagagtagagctgctcagcttctgcaagaggtatataagtcggaatgagttatca 1376  
OY 740 tgggtgagtagaagatagatcagagcttataatgtgcgcatatcagagcatatttgag 799  
DB 1377 gtattgtgagtagaagatagagcatataaagttaagagcatcttcgggaatattactgtg 1436

OY 800 tggttatgaagaagcctgtttgtgtcggagcaacatcagtaacatatacagatagagc 859  
DB 1437 tggttatgaagaagcctgt 1496  
OY 860 ggaatgtatgtactgtgtgaagagatgagagagcagaagaggaaggaaggaatgtcatgatt 919  
DB 1497 ggaatgtatgtactgtgtgaagagatgagagagcagaagaggaaggaaggaatgtcatgatt 1556  
OY 920 ttatcatgagttcttcaataataattatcatattgttgaagggaaggaaggaaggaagga 979  
DB 1557 ttatcatgagttcttcaataataattatcatattgttgaagggaaggaaggaaggaagga 1616  
OY 980 gtggaagggcagaacagaagagatgcgaattgtcgtgtgtgtgtgtgtgtgtgtgtgtgt 1039  
DB 1617 gtggtgtgacgaagaacagagatgagatgcgaattgtcgtgtgtgtgtgtgtgtgtgtgtgt 1676  
OY 1040 tgaattatgagtaggagtagcctgtcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1099  
DB 1677 tctgt 1736  
OY 1100 cacttgagaagggcagaacagaagagagagagagagagagagagagagagagagagagag 1159  
DB 1737 cacttgagaagggcagaacagaagagagagagagagagagagagagagagagagagagag 1796  
OY 1160 ttggaagagcagaattgt 1219  
DB 1797 ttggaagagcagaattgt 1856  
OY 1220 atgtggaacctaatgagcaaaagagagagagagagagagagagagagagagagagagag 1279  
DB 1857 atgtggaacctaatgagcaaaagagagagagagagagagagagagagagagagagagag 1916  
OY 1280 aaaaagctgtga-acagaagggagctcaatgacatattctatctgaagaagaacaaactca 1338  
DB 1917 gaaatgaatattgagtagaagaatgagcactgtgtgaatcacaagtgaaagtgagctgtg 1976  
OY 1339 ctccctgtgacctgtgtgaagagc-----atcaagtgagacttattgtgaagaagctg 1390  
DB 1977 aaatgtctccaaag 2036  
OY 1391 aggaatccaccgaatctcaagagatgaagctt----- 1422  
DB 2037 taactgacacacaaag 2096  
OY 1423 -cctgaagctctctatataaatttgaagttgaagaagcctgaagctgtgtgtgtgtgt 1480  
DB 2097 taactgacagttcctctgt 2156  
OY 1481 tctcgggagacattgt 1540  
DB 2157 tgt 2216  
OY 1541 ttgcaaaatataaaccattt---ggaataatgataaacaacacaaatgaagcatgtg 1597  
DB 2217 ttcaagagattatagagatcttaccgagatgagagagagagagagagagagagagagag 2276  
OY 1598 cagaatattatcagatagatagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1657  
DB 2277 gtaacatgtttctgtattgttcttagctcctgtgaattatttcttataacatttcc 2336  
OY 1658 tgcagagattatttgaagcag 1717  
DB 2337 tccagagccttacaatttggcagaagctgtggagagagagagagagagagagagagagag 2396  
OY 1718 cctcagaagcagatttatatagagatattgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1777  
DB 2397 ttctcagatccatgtgag 2456  
OY 1778 gaggcttgacacacaaatattgagcagatagatagagagagagagagagagagagagagag 1837  
DB 2457 gaggatgtgacacacagagctgtgcaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2516  
OY 1838 ggaattgt 1897

```

Db 2517 ggcctgcgtccttaaccgaatatagcaaatctcggagcagcattatatactctta 2576
Qy 1888 tatatgabatggagatgacatctcgtatctcgtatgtctccagtaacttgcggtagac 1957
Db 2577 tctatggttggcaatatacaacttaccctttagcaatttgaccatcattcgaatagacag 2636
Qy 1958 gaatgatgaaacccgacgaatgactgatttgcacaacaagaatagcaagaacttaagc 2017
Db 2637 gactgcttggaaatgaaatgctgtctggacaagcactggaagaatagaagaagcctagaag 2696
Qy 2018 atgcggaaagataagcaactggaagcttctggagaatataatagtaactatagttacaac 2077
Db 2697 gagcgggaaagattgctatacagaagcaccgaactccgaactcgtgttcttttggacc 2756
Qy 2078 gggaaagacccctcagcaaatgctatgaaagagatgcttcagaactcaacacagaataacct 2137
Db 2757 gggagcagaagtttgatatacatgatagcagaagtttgaagttacacatagagaataacct 2816
Qy 2138 cgaagaagaagcagaatttggaaagctgtatgataccttaagcattgccttatataatttg 2197
Db 2817 tgaagaaagacacacactcggggtctcaatttctatcaaccaggaagaatgataatttt 2876
Qy 2198 cctatgcagcaggttctcgatttggacacttaattcaagctggacgaatgacccag 2257
Db 2877 cctatgcgctgctgttccoggttctgtgctacttgggtgacaatgattcagaacttc 2936
Qy 2258 agggcagcttcaatgttcttacttgaatgacatagagataatggcaccgaaacagc 2317
Db 2937 aggaatgtcttcttggatctcagcattgtcttctgtccatggcaggggagcagttca 2996
Qy 2318 tctgttggctcttaatatcccaagcgaatccgagggctgagcattcgttgccttct 2377
Db 2997 gtcaatctgctcctgacatgccaagcgaagatcagcaccacagtcacatgata 3056
Qy 2378 tggaaagaagaacaatatagacagccgacgaagaaggaagaagcagacacatgtr 2437
Db 3057 ttgaaanaagccctctgattgacgatacagccctcagcctcgaagcacaataagtttg 3116
Qy 2438 aagggaatttagagttcgcgaagatcctcttctctatcatalgtgcgcagatgtttca 2497
Db 3117 aaggaaatgtgacatttaaltgaggtgctgttcaactatccactcgacagacatcccg 3176
Qy 2498 tctctcgtggacttaccctcagatattgagcgaggaagcagatagcatttggggagaga 2557
Db 3177 tgcctcaggggctgagccctcgaggtgaaagagccagcgtgcctcgtaggtaga 3236
Qy 2558 ggcgctgtgggaaagcaacttctgttcaactctgcaagagacttataagcccgtagac 2617
Db 3237 gtgctgtggggaagagcagatgttctcagctccttagagcgcttctatagcccttggctg 3296
Qy 2618 gacaaatcgttcttgaatgtgtgatactgacaagaattgaatgatactgctcgttccc 2677
Db 3297 gtccagtgtaattgatgtgacaagaagataaagcacctgaattgccagtggtcccgagac 3356
Qy 2678 aaatagcaatcgttctcctaagagcctgtgtcttcaactcgcagcagacttctgtgaacatcg 2737
Db 3357 accctggacatcgttctcctcagagccatcctgttgcactgcagcacttgcggaacatcg 3416
Qy 2738 cctatgtgtgacaacagcgtgtgtgctccttagatgatacgaagaagcgcgaatgacg 2797
Db 3417 cctatggagacaacagccgggtgcgtatcatcagaagaagatttgcaggaagcacaagag 3476
Qy 2798 caaatacatcttattatgaagctctccctgagaaatatacacaacacagttgactga 2857
Db 3477 ccaacatacacacttctcagacactcccgagaaataatacaccagagtagagagaca 3536
Qy 2858 aaggagacagcttctcgtcgcgcgaaacaaagactagctatttgaaggtctcttccc 2917
Db 3537 aaggaaaccaagctcctcgtgtgcgcgaaacagcagcatctcctcgtcgtctgtta 3596
Qy 2918 aaaaccccaaatattatgttggatgagcacttcaacccctgataatgacatgaga 2977

```

```

Db 3597 gacagcccatatatttgccttggatgaaagctacacagctcctgtagatacagaagtga 3656
Qy 2978 aggtgttcaagcatgccttggataaagccaggaaggaagacatgctatgttcaactc 3037
Db 3657 aggttcccaagaagccctggacaagcaggaagggccgcacctgcatgtgatacgccc 3716
Qy 3038 aacgctctcgtcaattcagaacagagattgatatggttcttcgacaattggaagataa 3097
Db 3717 accgcttgcacccatccgaatgagattatagttgtgttccaagaatgcaagta 3776
Qy 3098 aggaacaagaaactatcaagacccctcgtgagaatccgagacatatatttaagttatga 3157
Db 3777 agggagcatgacacatacaacagcgtcgtgcccaggaagaagcatctatttccatgta 3836
Qy 3158 atgcacagtcag 3169
Db 3837 gtgtccaggtc 3848

RESULT 2
AAd03504
ID AAd03504 standard; cDNA; 4279 BP.
XX
AC AAd03504;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype A) protein"
FT allele
FT /tag= b
XX
PN W020123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX WPI: 2001-235373/24.
XX P-PSDB: AAE00308.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 9; Page 85-90; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein, PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP

```



QY	2138	cgaaagaagcaacagattatttggaagctgltatgatctacagctgccttaataatttg	2197
Db	2817	tgaaagaagcaacaactctcggggtcccatcttctacaccaggaagaagatgatttct	2876
QY	2198	ccctctgacagaaggttctcgaattctgagccctatttaactaaagctcgagacgattccag	2257
Db	2877	ccctactgcgtctgtctccggttttgggtccactgtgtgcgaatgaatgattcaagaacttcc	2938
QY	2258	agggcatgttcaatgatttcttaactgcgaattcgatatafgagctatvgccalcgnaaaacgc	2317
Db	2937	aggaatgtctcttcttgatacttcacagcatgtgtcttctgtgcacatgycagtgycagtgca	2996
QY	2318	tcgtcttgcgcctccggaatatcccaagggccaaatccggggctbctgcatactgtctgcctgt	2377
Db	2997	gttcaattgtctctctgactaagcccaagaagccaagttalcagtcagcccaagccacatcagtaca	3056
QY	2378	tgnaaaagaanaaccaatatalagacagccgcagttaaagaaggaanaaaagccagacacatgtg	2437
Db	3057	ttgnaaaagaagccctctgattgtacagcttaagccctccacggctccagcccaataatagcttg	3116
QY	2438	aagggaattatagagttctcgagaagctcttcttctataccatgtgcgccagaatgtttca	2497
Db	3117	aaggaaatgtgacacttaatcagagtgctgttctcaactatccactcgacagacacatcccg	3176
QY	2498	tcgccctgtgcttataccctccatgatcgaagcgaagaaagcaagtaagatttgggggagca	2557
Db	3177	tgccccagggggctcgagcctctcgagtgaaagaagggccagacgcgtgcctctgaagtacga	3236
QY	2558	ggcgctgtggaagaagcaactctgtctcaactctctgcagagacttatagaccctgtgcag	2617
Db	3237	gtgccttcttggaagagacagactgtgtccactccttaagagcgcttcatagccccctgtgtg	3296
QY	2618	gacaagcgccgtttgtatgagtggtgtagcgcgaagaatctgaagttaagtggtccctctcc	2677
Db	3297	gttcaatgtcctcaattgtatggtgcgaagaagataaagcaactctgaatgttccaatgtgtcccgagac	3356
QY	2678	aaatagcaatcgttctccccaagagcctgtgtgcctctcaacgcgcagcatctgtctgnaaacatcg	2737
Db	3357	accgtgcatctgtctctcagaagcccatctcctgtttgactgcagacttgcgcgagaacttgg	3418
QY	2738	ccctctgtgacacaagcgcgtgtgtgtgcatttatgatagtatcacaagaagccgcgaataatcag	2797
Db	3417	ccctctgtgagacaacaagccgcgggtcgtatcacaatgaagaagatttatgagaagccgaagaggg	3476
QY	2798	caaatcaccaactcttcttatgaaagcgctccctcgatagaatatacaacacaagctgtgacga	2857
Db	3477	cccaactacaccacttcatctgagaacactctccctgagagaataacacaacacagatgaaagaca	3536
QY	2858	aaggagcacagccttctgtgcggtccagaanaacaaagactagctatctgcaaggtgccttctcc	2917
Db	3537	aaggaacaccgctctctctgtgcgcagaanaacagcgtcatctgcatagctgcgctctgttga	3596
QY	2918	aaaaaaccaaatatttattgtttgtagtgggcacacttccacccctgataatgacagtatga	2977
Db	3557	gacagcctcatatttgcgttcttgtagaagcttacaacagctcttcgataaagaagaatgaa	3656
QY	2978	aggtgtgttcagcacatgccttgataaagccagagccaggaaggaacatgactgtgacatc	3037
Db	3657	aggtgtcccaagaagccctgtgacaaagccagagaaggtccgcacctgtcatgtgacgtcc	3716
QY	3038	acagagcctctgcattctcagaagaagcagatttgatagtggttcttgcacaatgtgaagataa	3097
Db	3717	accgctgttccacatctcagaatgcagatattaatggtgtgttcaagaatgtgcaaaagta	3776
QY	3098	aggaacaagaagacatcacaagacgtcctcgtaagaatcgaaacatatatttaagtattgga	3157
Db	3777	agggacatgtgcacacatcaacacagctgtgtgcgcagaagaagcatctatatttccatgtgtca	3836
QY	3158	atgcacagctcaag	3169
Db	3837	gtgtccagagctg	3848

RESULT 3  
 AAD03505  
 ID AAD03505 standard; CDNA; 4279 BP.  
 XX  
 AC AAD03505;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.  
 XX  
 KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KM MDRL; drug bioavailability; transgenic animal; genetic model; ss.  
 XX  
 OS Canis familiaris.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 17..3862  
 FT /\*tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (genotype B) protein"  
 FT replace (91, T)  
 FT /\*tag= b  
 FT allele  
 FT replace (607, C)  
 FT /\*tag= c  
 XX  
 PN WO200123540-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26767.  
 XX  
 PR 28-SEP-1999; 99US-0156510.  
 XX  
 PA (GENT-) GENTEST CORP.  
 XX  
 PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Pattem CJ;  
 DR MPI: 2001-235373/24.  
 XX  
 P-PSDB: AAE00309.  
 XX  
 PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX  
 PS  
 XX  
 PS Claim 9; Page 93-99; 11pp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDRL) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein, PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 XX  
 SQ Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Db 843 ttgaagagtcacacaaaatttagaagaagctaaaggaattggagataaagaagctatca 902  
QY 260 ctccaaaagtgctctctgtggtcgtgtgactctttagaattggaactaagacttgct 319  
Db 903 cggccaacatctcatgtggtccgcttcttattgatactatgacatactgcttgct 962  
QY 320 ttgtgataagacctcccttcttattccttaattgagaacctgatatcatcgagcttcc 379  
Db 963 tctggtatggagacctcctctgtctctccag-----tgatatctcattggacaagtac 1016  
QY 380 tgcgtcttcttcttaagtgtaatccatagcagttatgcatctggagcagctccctac 439  
Db 1017 tcactgctctcttctctgttaattaaatggtctttagtattggaacagcactcccaagca 1076  
QY 440 ttgaaccttcgcaatagcccgagagactgcttcaatttccagagttattgataga 499  
Db 1077 ttgaagatttgcaaacgcaagagagcagcttatgaaatcttcaagataattgacata 1136  
QY 500 aaccagtatagataactttccacagctgataataaacctgaaatccatagaagaactg 559  
Db 1137 aaccagcatctgacagctattcgaaaggtggaactaaacagataataataagggaaatt 1196  
QY 560 tgaatttaaaaatgttcttcaatatccatcaagaacacatcatcaagaattctgaag 619  
Db 1197 tggattcaaaaatgttccacttccagttaccccttcgaaaagaagttaaagctttaaag 1256  
QY 620 gtccgaatcccaagaataagctctggaagacagtcgcttgctgctcaatgagcgtg 679  
Db 1257 gtcccaacctgaaggttcagagtgggcagacagtggtggtggaacagtggtgcg 1316  
QY 680 ggaagatcacgtgagttccagctctgcagaggtatatagtaccggatagctgttaca 739  
Db 1317 ggaagagcaacagaccgtgcagctgagtgagaggtcctatgacccacagatgagcgtgct 1376  
QY 740 tgggtgagtgagaatgatacagagctttaaattgtgcggcatatccagacacattggag 799  
Db 1377 gtattgtgagcagagacatagacccaataagcactcctcgggaaattactcgtg 1436  
QY 800 tgggtatgcaagagccgtgttctgtcggagaccatcagtaacataatcaagattgac 859  
Db 1437 tgggtgagtcagagccgtgtgttctgtccacacagataagctgaaacacttgcataggcc 1496  
QY 860 gagaatgtgtaactgataagaatggaagagcagcaagggaaacatgctgatatg 919  
Db 1497 ggcgaataatgtaacatgataagattgagaatggaagctgttaagaagacatgctcattc 1556  
QY 920 ttatcatggagtttccctataataatttaacatgtgtaggggaaaaagagctccaatga 979  
Db 1557 ttatcatgaaactacactataaatttgacactcgtgtgagagagagggcccaagctga 1616  
QY 980 gtggaagggcaaaaacagagagctgcaattgtctgtccttgaatcccaagcttcc 1039  
Db 1617 gtgggtggaacagaaacaggaatcgccatctgtcgggcccgtgttcgcaaccccaagatc 1676  
QY 1040 tgattttagatgagactagctgcctccctggatctcagaagaagctcagctgtcaagctg 1099  
Db 1677 ttctgcttgagatgagagcaagctcagctcctggacactggaagaggaagcagtggttcaggtg 1756  
QY 1100 cacttggagaagcgagcaaaagtcgagactacaatcgttgtagacaacacgacttctacta 1159  
Db 1737 ccttgataagcgcaagaagaagccgagactacatgtgatagctcatcgtttgtctacag 1736  
QY 1160 ttcgaaagcgagatttgatttgaccctaaagagatggaatgctgcggaagaagaagac 1219  
Db 1797 ttctgaatggccgagttctcgtgttttgatgtagtgagatcatgttgtagaagaagaaatc 1856  
QY 1220 atgctgaactaagtcgaacaaagaggtctatatcttaccctgtgtagcacaagatatca 1279  
Db 1857 atgataacatcagtaagaagaagaggaattacttcaaaactgttcaacaatgtagacaagaag 1916  
QY 1280 aaaaagctgatga--acagatggagtcaatgacataattctactgaaagaagaaccaatca 1338

Db 1917 gaaatgaanaattgagtagaanaatgccaactggtgataatccaaaagtgaagtgtgctgg 1976  
QY 1339 ctccctctgactctgtgaagagc-----atcaagtcaagactcatttgacaagctg 1390  
Db 1977 aaatgtctccaaaagaatctcaaggttccaggtttaataaagaagaatcaactcgaagagta 2036  
QY 1391 aggaatccaccacatctaaagataagctt----- 1422  
Db 2037 tactatgaccacaaagccaaagacagaagaagcttggtaacaaagagacttgaatgagatg 2096  
QY 1423 --cctgaagctctctatthaanaattttaaagtttaacaagcctgaatgagcttgg 1480  
Db 2097 tactccagatttccctctggaagatctggaagctggaactgaatgagcttatttg 2156  
QY 1481 tctctgggaacattgagcttccgtcttaaatgaaactgttcaatccagatatttccatcat 1540  
Db 2157 tggttgtgatatatttggtcattataataacgagagcctgcaacagcatthaataat 2216  
QY 1541 ttgcaaaaattaaacatgtt---ggaataatgataaaaccacataaagcatgag 1597  
Db 2217 ttcaagatattagagatcttaccgagatgagatccctgaaacaaacgacagata 2276  
QY 1598 cagaattattccatgataatttgcatttgggtgttatgttctgtgcagtattcc 1657  
Db 2277 gtacatgttctctgtatgttcttctagctccttggaatcatcttcatatcaatttcc 2336  
QY 1658 tgcagagattattttaacgagagcaagggaaattttaaagcatgagattagaacttgg 1717  
Db 2337 tccagagcttcacatattggaagaagctggggagatccctacataagcgttcgtatag 2296  
QY 1718 ccttcaaaagcatttatatcagaagatatgtcgtggttgaatgaaagaagaacagcag 1777  
Db 2397 ttctcagatcatagctgagacagatgtcagctggtttgtagtccctaaacacactg 2456  
QY 1778 gagcttgaaacaaatatttgccaatagatagacaaatctcagaagggacagagcttca 1837  
Db 2457 gagcatgacaacaaagcttgccaatgagtcggtcccaagttaaagggcctatggttcca 2516  
QY 1838 ggaattgctcttaacacaaaatgccaactaacatgagacttcaagtatcatcttcc 1897  
Db 2517 ggtcttgctgcatatcaccagaatagcaaatcttggagacagcatatatatccctaa 2576  
QY 1898 tatatgagtcggagatgacatctcgtatctgagtatgtctccagatcttgcggtgacag 1957  
Db 2577 tctaigtgtgcaaatcaacacttactccttaagaaatttgtaaccatattgacataagcag 2636  
QY 1958 gaatgattgaaacccgagcaatgactgatttggccaacaaagataagcaagaactaagc 2017  
Db 2637 gagttgttgaatgaaaaatgtgtctggaacaaagcagcaagaagaaagagctagaag 2696  
QY 2018 atgctggaagatagcaactgaaagcttggagaatatacgttactatagtgltcaatlaacaa 2077  
Db 2697 gagcttgggaagatgtcacaagaagccatcgaaacacttcogaactgtgttcttggctc 2756  
QY 2078 ggggaagaagcccttcgagcaaatgatatgaagagatgcttcgaactcaacacagaatcc 2137  
Db 2757 ggggacgaagtttgaatacagtatgacagagatttgcaagttccataagaacacttc 2816  
QY 2138 cgaagaagaacacagatatttgaagagctgttatgatactcagccagcttataatttg 2197  
Db 2817 tgaagaaagcacacatcttcggggcttccatatttccataccacaggaacagatgataattc 2876  
QY 2198 cctatgacagcaggttccgatttgaagcctaatccaagctggaagctggaatgacccag 2257  
Db 2877 cctatgtcgtgctgttccoggttgggtcctactcgtggcgaatgagatgttcaacttcc 2936  
QY 2258 aggcagatgataagttttaaactgcaatgtcataatggaagctatgacatcggaacagc 2317  
Db 2937 aggaatgtcttcttggatctccagctattgtcttggltgcocatgagtggggagagltca 2996  
QY 2318 tcgtttggctcccgataattccaaagcgaatctgggggctcgtgcagctgttctgctgt 2377  
Db 2997 gtcatattgctcctgactatgccaagccaagatatacagcagccacgtcatcatgata 3056

QY	2378	tggnaaagaaacccaatatgaaagcccgagtcgaagaagggaagaaagccagacacatctg	24377
Db	3057	ttgaaataaagccctctgaatttgacagcttacaagccctcacaagccataaataagttctg	31166
QY	2438	aaggaataattagagttctcgaaagctctctctctctcatccatgctcgccagatgtttcca	24977
Db	3117	aaggaaatctggaacttaataagagtcgtgttcaactatccctcgaccagacatcccg	31766
QY	2498	tcctccgtggtctatccctcagatctgaagccgagaaagacaagtaagcatctgtggnagga	25577
Db	3177	tgctccaggggcttgagcctctcgaggtgaagaaggccagacgctgtgcccctcgaggtacga	32366
QY	2558	ggcgctgtgggaaagaagcactctgtccaactctctgaaagagatctatgacccctgtgcaag	26177
Db	3237	gtggctgtgggaaagagacacagctgtctcaactcctctaaagcgctctctacagccctctggtg	32966
QY	2618	gacaaagcgctgttgatcgtgtgtgaatgcacaagaagaaattgaagttaacgtgtgctccctcc	26777
Db	3297	gttcagtgctaatgtatgtggaagagagataaagacacctgaaatgtccagtggtctccgaagac	33566
QY	2678	aaataagcaatcgtctctccaagagccctgtgctcttcaactcgacagcatctgtcgaagacatcg	27377
Db	3357	accgtggcactcgtgtctctcaagagcccaactcctgtttacgcgcagcatctgcggaaacattg	34166
QY	2738	ccctctgttgacaagaagcgcgtgtgtgtgcattagaatgaatcacaagaagccgcaaatctgacg	27977
Db	3417	ccctatggaagacacaagccggtctcgatcaacatgaagaagatctatgcaggaagccaaaggag	34766
QY	2798	caaatatccattcctttatgttaaaggtctccctctgagaaatatacaacacaagttgtgacgtga	28577
Db	3477	ccaacatatacccaacttcatctcgaagacactccctcctgagaataatacaaccagatgagagaca	35366
QY	2858	aaggagaacaggtctctctgtggggcccaagaacaaagactaactatgtgcaagggtctctccc	29177
Db	3537	aaggaacccagctctctctgtggcccaagaacagcgcatctgcataagctcgcgtctgttga	35966
QY	2918	aaaaacccaataatttatgtgtgatagagccacttcagacccctcgataatgatagtacgtgga	29777
Db	3597	gaacgcctcatatttctgtcttctgtgatagaagctatacagactctgtgatatagaagaagtga	36566
QY	2978	aggtgtgttcagcagtgccctctgataaagccagagccaggaacggaagacatgctagtgtgctac	30377
Db	3657	aggtgtgtccaagaagccctctgacacaagccagaggaagggcccgcaacctgtgcatgtgacgcc	37166
QY	3038	acaaagctctctgcgaattctcaaaagccagatcttgatgtgttctgtcacaatgtgaaagataa	30977
Db	3717	accgctgttccacacatcccaatcacaatgcagatcttaataagttgtgttccagaaatgtgcaaaag	37766
QY	3098	aggaacaagagacatccatcaagaagcctcctgtgaagatcgagacataatcttaagtattgacga	31577
Db	3777	agggagcatgtgacacatcaacaacagctgtgtgcggccagaaagacatcatattttctccatgtg	38366
QY	3158	atgcacagctcag	3169
Db	3837	gtgtccagagctg	3848

RESULT	4
AAD03506	
ID	AAD03506 standard; cDNA; 4279 BP.
XX	
AC	
XX	AAD03506;
DT	13-JUN-2001 (first entry)
XX	
DE	Dog P-glycoprotein (Pgp) allelic variant (Genotype D) cDNA.
XX	
KW	Dog: P-glycoprotein allelic variant; Pgp: multidrug transporter;
XX	MDR1; drug bioavailability; transgenic animal; genetic model; ss
XX	Canis familiaris.
XX	

FH	Key	Location/Qualifiers
FT	CDS	17..3862
FT		/*tag= a
FT		/product= "Dog P-glycoprotein (PGP) allelic variant
FT		(genotype D) protein"
FT	allele	replace (91, T)
FT		/*tag= b
FT	allele	replace (607, C)
FT		/*tag= c
FT	allele	replace (1001, T)
FT		/*tag= c
FT	allele	replace (3458, A)
FT		/*tag= c
XX		
PN	WO200123540-A2.	
PD	05-APR-2001.	
XX		
PR	28-SEP-2000; 2000OWO-US26767.	
XX		
PR	28-SEP-1999; 99US-0156510.	
XX		
PA	(GENT-) GENTEST CORP.	
XX		
PI	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;	
XX		
DR	WPI; 2001-235373/24.	
XX	P-PsDB; AAE00310.	
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful	
PT	for determining the bioavailability of drugs and for screening for dog	
PT	PGP inhibitors -	
PS	Claim 9; Page 102-107; 11pp; English.	
XX		
CC	The invention relates to dog P-glycoprotein (PGP) also referred	
CC	as multidrug transporter (MDR1) and nucleic acids encoding them,	
CC	The invention also includes fragments and biologically functional	
CC	variants of dog P-glycoprotein. PGP and their nucleic acids are	
CC	useful for determining the bioavailability of drugs and for	
CC	screening PGP inhibitors. They are useful for the diagnosis and	
CC	treatment of conditions characterised by PGP activity, by	
CC	reducing or increasing PGP activity in a cell. PGP nucleic acids	
CC	are used as oligonucleotide probes. Complements of PGP nucleic	
CC	acids are useful as antisense oligonucleotides, to induce a PGP	
CC	'knockout' phenotype. They are used to prepare a non-human	
CC	transgenic animal, which are valuable as genetic models for	
CC	human diseases.	
CC	The present sequence is dog P-glycoprotein (PGP) allelic variant	
CC	(Genotype D) cDNA. The PGP enzyme functions as an efflux pump	
CC	exporting small molecules across the cell membrane. This enzyme	
CC	is a member of the ABC transporter family.	
XX		
SQ	Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;	
<hr/>		
	Query Match	35.6%; Score 1130; DB 22; Length 4279;
	Best Local Similarity	62.4%; Pred. No. 1.4e-295;
	Matches 1879; Conservative	0; Mismatches 1085; Indels 48; Gaps
QY	200 ttgtaaagtgatacacagaatctcaagaatgacaagaatttggcattaagaagtactatag	259
Db	843 ttgaagaagtcacaacaaaatttagaagaagcctaagaatttggatagaagaagtacata	902
QY	260 cttaacaagtgctcttgtgtgtgtgtgtaccttttttgaatgaaggaccatttgactgctt	319
Db	903 cgagcaaacattctcatctgtgtgcgctttcttatttgatcagcatcataatcgctctgctt	962
QY	320 ttgttgatgaacccctctgatctcttaatgaagggaaccttgataccaacgggaactgttc	379
Db	963 tctgttatvgagaccctctgtcctctccag-----tgaatatcatttggacaagtac	1016
QY	380 ttctgtttcttctttagtgaatccabagcagttatgtgcatvtggaagcagctccctact	439

Db 1017 tcaatgtcttcttcgtatataattggyctttagattggaagcaccaccaagca 1076  
QY 440 ttgaaaccttgcgaatgcccgaagagctgtccttcaatttccagttttataga 499  
Db 1077 ttgaagctttgcgaacgaagagagagctttagtaatttcaagaatttcaatla 1136  
QY 500 aaccagatataataacttctccacagcttgatataaaacctgaatccataagaagactg 559  
Db 1137 aaccaagatgtacacgctattccgaagagtgacataaaccgataatataaggaaatt 1196  
QY 560 tgaataaataaagtcttcttcaattatccataagaaccatactcaagctctgaag 619  
Db 1197 tgaatttcaaaaatgttcaacttccattccctctcgaaaagaagttaagattcaaaag 1256  
QY 620 gtctgaattcgaattaagcttggagagacagtcgcttggctgcgttcaattggcagt 679  
Db 1257 gtcccaacctgaagtttcaaggtggcgacagatggcgctgttgggaacagtgctgcg 1316  
QY 680 ggaagagtaacggttagtcacagcttccgagaggttatatgaaccgagatgattatca 739  
Db 1317 ggaagagacagacccgtgcagctgtgagtgccctatagaccacagatgycatgctc 1376  
QY 740 tggtagatgagaatgacatcagagcttbaaattgtgcgcatatcagagacataittgag 799  
Db 1377 gtatgtatgagcagagacattagaccataaattgaagcattcttcgggaattactggtg 1436  
QY 800 tggtagtcaagaagcgtgttcttgcgggaccacacatcagtaacaatcaagtatgac 859  
Db 1437 tggtagatgagaagacgttgttgccttgcacacagatagctggaaaaacttgcctatggcc 1496  
QY 860 gagaatgagtacgtaacgataagatgagagacagcaagggaagcaaatgcgtatgatt 919  
Db 1497 ggcgaataatgtacacatgataagattgagaagagctgttaagaaagccaatgtccattg 1556  
QY 920 ttatcatgaggtttccataataattatatacatgtgtgtaggggaaaaagagctcaatga 979  
Db 1557 ttatcatgaaactacactataaatttgcacactgtgttggagagagagggccacgtga 1616  
QY 980 gtggagggcgaaaaacagagatgcgcaattgtcgttcgttcgttcgaaaccccaagatc 1039  
Db 1617 gtggtagcagaaaaacaggaatccgcatctgcgggcccgtgttcgcgaaccccaagatcc 1676  
QY 1040 tgaatttagatgagctacgtctgccttgcgaattcagaagaacagtcagctgttcaagctg 1099  
Db 1677 ttctgcgtgagtgaagcaagctcagctctgcgacacagaaagagagcagtggttcaggtg 1736  
QY 1100 cactggagaagcgagcaaaagtcggaactacaatcggtgtgtagcacacccgactttacta 1159  
Db 1737 cccgtgataaagcgagaaaaagccggaactacatgttatagctcatcgttgcctacag 1796  
QY 1160 ttccaagatcgaattgttgaaccctaaagagtggaatgctgicggagagaagagacac 1219  
Db 1797 ttctgaatgcgagatgtctgtcgtgttgaatgagagtcattgtgttggaaagaaatc 1856  
QY 1220 atgtctgaactaatgccaagaaggtctatatatctacactgtgtgtagccaagatala 1279  
Db 1857 atgatagaactatgaagaagagggcatttacttcaaacctgttcaaatgtcagaacagag 1916  
QY 1280 aaaaagctgatga-acagatgagatcaatgacataattctactgaaagaagaaccaatca 1338  
Db 1917 gaatatgaatgtagttgaanaatgcacactggtgaatccaagaagtgaagtgtccttg 1976  
QY 1339 ctctcctgcagctctgtgaaagc-----atcaagtcagacttctatigaacagctg 1390  
Db 1977 aaatgtctcccaaaagatccagggtccagtttaataaaaagaagatacaatcgcaggtga 2036  
QY 1391 aggaatccaccaatctaaagataagttct----- 1422  
Db 2037 taactgacccaagagcgcaagagaaagctgtgtacaaaagagacttgatagaaatg 2096  
QY 1423 --cctgaagctctctatataaaattttaagtttaaacaaagcctgaatgaccttgtgg 1480

Db 2097 taactccagttctctctgtagagattctgaagctgaactcaactgaatgacctatttg 2156  
QY 1481 ttctgggagacatgctctgttcttaatggaactgttcatccagatatttccatctc 1540  
Db 2157 tggtagtaataatttctgtctataaacygagccgcgaacccagcattttcaataat 2216  
QY 1541 ttgcanaaatatataccatgtt--ggaataatgtataaaccacataaagatgatg 1597  
Db 2217 ttcaagatataatagagcttccatccgagatgagatccctgaaacaaaacgacagata 2276  
QY 1598 cagaattatccatgatatactgcattttgggtgtatttgccttgcgtatttca 1657  
Db 2277 gtaacatgttctgtatgttcttctagtccttggaaatatttcttcttattccattcc 2336  
QY 1658 tgcagggatattttacgagcagaggaaggaatttaacgatgatattagacctgg 1717  
Db 2337 tccagggcttcaacttggcaaaagctgggagatccctcaacagcgcttcgaataatg 2396  
QY 1718 ccttcaaaagccatgttatatagataltgctgtgttgaataaagaaaacagcacag 1777  
Db 2397 ttctcagatccatgtcgaagacagatgtcagctgtttagatgcccttaaaacacacag 2456  
QY 1778 gaggcttgacaacaatlatagccatagatlagacaacaattcaaggagcaacagttcca 1837  
Db 2457 gagcattgacacaacagccttgccaatgagtgcgctcaagttaaaagggtatagttcca 2516  
QY 1838 ggaattggcgtcttacaacaaaatgcaactaaatgggaacttccgttatcccttca 1897  
Db 2517 ggcctgtcgtlatatccacagatatagaaatcttggagacagcattatlatccctaa 2576  
QY 1898 tatatgagtagagatgacatccctgattcgaatattgtgcacagactactgcgtgacag 1957  
Db 2577 tctatggttgcgaattacatcaacttacttactcgttagcaattgtgccatcatgtacag 2636  
QY 1958 gaatgattgaaacccgcagcaatgacatgatttgcacaagaataagacaagacttaagc 2017  
Db 2637 gagtgtgtgaatgtaaaatgttcttgagacaacagctgaagaaataaagaagataag 2696  
QY 2018 atgctggaagaatgacactgaactgttggagaatatacgtacatagttgcatlaacaa 2077  
Db 2697 gagcttggaaagatgtgtacagaagccatcgaaaacttccgaacactggttcttcttgaac 2736  
QY 2078 gggaaaagaccttgcgagcaaatgtatgagaagatgcttcaagactcaacacagaataact 2137  
Db 2757 gggagcagaagtttgaatatagatgacagagtttcaagtaaccatacagaataact 2816  
QY 2138 cgaagaagaagcaaatatttgaagctgttatgcatcagccaagccttatattttg 2197  
Db 2817 tgaagaaagcacacatcttcgggggtlctcatlctatcaaccagcaatgtatgtatttt 2876  
QY 2198 cctatgcagcagggtttgcgaattggaagccttaattcaagctggagcgaatgacccag 2257  
Db 2877 cctatgctgcgtgttcccggttctgtgtcctactcgtgtgagcaaatgagatltcaacttc 2936  
QY 2258 agggcatgtlcatagtttttacttgcgaattgcatatgagctatgccaatgcgaaaacgc 2317  
Db 2937 aggaatgttcttcttgatctccagctatgtcttgtgccaatgacagtgggagagttca 2996  
QY 2318 tctgtttggctcctgaaatctccaaagccaaatccgggggctgcgcacatcgtttcctgt 2377  
Db 2997 gtcatattgcctcctgactatgcgaagccaaagatacagccagccacgtctcatgtatca 3056  
QY 2378 tggaaaagaacccaatatagacagccgagctcaagaagggaagaaaacgacagacagtg 2437  
Db 3057 ttgaaaagaacccctctgattgacagctacagccctcaagcgtcctaagcaataatcglttg 3116  
QY 2438 aaggaatttagagtttgcgaagaatctcttctctatccatgtgcgccaagatgttttca 2497  
Db 3117 aaggaataatgacatttaatagatgtgtgtcaactatccacctgcagaccagaatcccg 3176  
QY 2498 tctccgtggtctatccctcaagtatgtagcggaagaaagacagatgtattgtgggagca 2557  
Db 3177 tgtcccaagggtctgagcctcagagtgaaagagggccacagacgtgcctcctglaagtagca 3236



QY	2558	gcccctctgggaaagaaagacttcggttcaactctctcgaaggaatttataccgcgtgcaag	2617
Db	3237	gtggcgtctgggaagagacacggttcttcagctctctcgaagggctctctatgaccccttgctg	3236
QY	2618	gacaagtcgtctcttgatgvtgtgtaatgacaaagaattgaatgttaacagtgcgtccgttccc	2677
Db	3297	gttcacgtcctaatttgatgvtgaaagaagataaagcacctgaatgattccagtgctccgagac	3356
QY	2678	aaataagcaatccgttctctcgaagagcccggtgctcttcaactgtcagacattgctgaacacgtg	2737
Db	3357	accggtgacatctggtctccaggaagcccaatccctgtttgacgtgcagcattgcccagaacaattg	3416
QY	2738	ccctatgtgacacaagcccggtgtgtgtccattagatgagatcaaaagaagccgcaaatgag	2797
Db	3417	ccctatgtgagacaacacagccggtgtcgtatcacatgaagaagattgtgcagagcccaagaaggtg	3476
QY	2798	caaatatcatctcttattgtaaagtgctccctgtgagaataacacacacaagttgactga	2857
Db	3477	ccaacatcaccccaacttcatctgagagacatccctcgagaaatatacacaccagatgagagca	3536
QY	2858	aaggaagacacagcttctctgctgagcgaagaaacaaagactgataattgcaagggtcttctccc	2917
Db	3337	aaggaacaccagctctctgtgtggtccagaacaaacagcagattgccaatgactgcgccttgta	3596
QY	2918	aaaaaaccaaatattatgttgtatgtgaggtccacttcacccctcgataatgacagttgaga	2977
Db	3357	gacagccctcatatttgcgtcttggatgagaagcctacatcagcctctgtatcagaagaagtga	3656
QY	2978	aggtgtgttcagcatgcacctgtataaagccaagcaggaagcggaaagacatgctctagtgtcactc	3037
Db	3657	aggtgtgtccaagagccctgtgacacaagccaggaaggtccgcacccgtgacttgatcgccc	3716
QY	3038	acagagcctctgcaattccgaatcgaagacgcagattgtatgtgtgtctgcacatggaagataa	3097
Db	3717	accgctgttccaccatctcgaatcagaatgacattatagttgttctagaagtgcgaagta	3776
QY	3098	aggaacaaagactctcatcaagagcctctctgaagaatccgaacatcatatttaagttagtga	3157
Db	3777	aggaagcatgtgcacacatcaacacagctgtgcggcccaagaagaatcatatttccatgtgta	3836
QY	3158	atgcacacagtcag	3169
Db	3837	gtgtccagcgtg	3848
RESULT 5			
AAD03489			
ID	AAD03489 standard; cDNA: 4317 BP.		
AC			
XX	AAD03489:		
XX			
DT	13-JUN-2001 (first entry)		
XX			
DE	Dog P-glycoprotein (Pgp) cDNA #2.		
XX			
KM	Dog: P-glycoprotein; PGP: multidrug transporter; MDR1;		
KW	drug bioavailability; transgenic animal; genetic model; ss.		
XX			
OS	Canis familiaris.		
XX			
FT	Key		
FT	CDS		
FT	70..3912		
FT	/tag="a		
FT	/product="Dog P-glycoprotein (Pgp) #2"		
XX			
PN	WO200123540-A2.		
XX			
PD	05-APR-2001.		
XX			
XX	28-SEP-2000; 2000MO-US26767.		
XX			
XX	28-SEP-1999; 99US-0156510.		
XX			

PA	(GENT-) GENTEST CORP.
XX	
PI	Stocker PJ, Steime1-crespi DT, Crespi CL, Relf TC, Patten CJ;
DR	WPI; 2001-235373/24.
DR	P-PSDB; AAE00304.
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful for determining the bioavailability of drugs and for screening for dog PGP inhibitors -
PS	Claim 1; Page 66-72; 11np; English.
CC	The invention relates to dog P-glycoprotein (PGP) also referred as multidrug transporter (MDR), and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog P-glycoprotein. PGP and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterised by PGP activity, by reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.
CC	The present sequence is dog P-glycoprotein (PGP) cDNA. The PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.
SQ	Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;
Query Match	35.5%; Score 1126.8; DB 22; Length 4317;
Best Local Similarity	62.3%; Pred. No. 1e-294;
Matches 1877; Conservative	0; Mismatches 1087; Indels 48; Gaps
OY	200 ttgtaatgatatcacgaattccaaatcgaagatggatttgcataaaaggactatag 259
Db	893 ttgaagaagtacacaaaaatttagaagaagctcaagaatttggataaagaagctaca 952
OY	260 ctccaanaagtctctcttggtgctgtygtacttccttaatgaagaacctatgtaactgc 319
Db	953 cgsgcaaacattctcatgttgcgcgtttcttatgatcatgatcataatgctctgcgtt 1012
OY	320 ttgtgtatgaagacccttcgtgatctttaatgaggaaccttggatatcacggagactgc 379
Db	1013 tcttgtatgaggacccttcgtgcccctccag-----tgaatatcatattgacaggtac 1066
OY	380 ttcgcgtttctcttagtgaatccatcagatgaattgcatitggagcagagctccctact 439
Db	1067 tcaactgtctcttcttcttgatattaatttggcgcttttagatttggacagagatcccaagca 1126
OY	440 ttgaaccttcgcataagcccgagagagctgcttccatatatttccagtgattatagaaga 499
Db	1127 ttgaagcatcttgcacaagcagaagagagcagcttatgaaatctccaagataattgaaata 1186
OY	500 aaccagatitgataacttccattccacagcggatataaaccttaalccatagaagaagcty 559
Db	1187 aaccaagcatigacgctctatcgaagagtgagacataaacacagataataltaagggaatt 1246
OY	560 tggaaatttaaaaagtcttcttcaattatccatcaagaaccatctataccaagatttgaag 619
Db	1247 tggaaattcaaaaattgtcacttcagttacccttcttcgaagaagttaaagattctaag 1306
OY	620 gctctaattcgaattaaagcttgagagagagagctgccttggttcggtccatgtgcaag 679
Db	1307 gctccaacctggaaggttcagagtgggcagacagtgcgcttggttggaaacaagtgcctg 1366
OY	680 ggaagaagtagcttagtccagcttcgcagaagattatattatccggaatggatgcttatcaa 739



Db 1367 ggaagacagaccgtgcaagctgacgtcagagcgtcctatgaccccaagatgacgtcct 1426  
Qy 740 tgggtggtggaatgacatgaagccttaattgctgggcatctacgagacatcttgag 799  
Db 1427 gtactgctggacaggaactttagacccaataatgaaagcactcttcggyaaattctggtg 1486  
Qy 800 tggtaatcaagaccctgttctgctgggacaccatcaagtaacatataagatctgac 859  
Db 1487 tgggtgagtcaggaagccctgtgttctgcccacagatagcggaaacattctgcattggcc 1546  
Qy 860 gagaatgtgacatgatagaatgagagagcagcagaaggaagcaaatgctatgatt 919  
Db 1547 gcgaaatgtccaccatgcatgattgaaagcgtttaagaaagccaatgctcattagtt 1606  
Qy 920 ttatcatgaggtttccctaataattataattgtgtaggggaaagagagctcaaatga 979  
Db 1607 ttacacagaaactcaactaataattgacactcgtgtgagagaaagagggccggcgca 1666  
Qy 980 gtgagggcagaaacagagaatcgcaattgctgctgaccttagctgaaaccccaagattc 1039  
Db 1667 gtggtgagcagaagaacagagatcgcaattgctgggcccgtgttcgcaacccaagattc 1726  
Qy 1040 tgattttagatgagcctacgtctgctccttgatctcaagaagcaagtcaagctgttcaagctg 1099  
Db 1727 tctgctgagatgagcagcactgcaactctggaactgaaagtgaagcagtggttcaagtg 1786  
Qy 1100 cactggaagagcgagcagcagaagctcggaactacatcggtgtagcacaacggattctacta 1159  
Db 1787 cccgtgataagcgccagaaagggccggactacatctgtagatgcacatcggttgcctacag 1846  
Qy 1160 ttcgaagtgcagatttgattgacccctaagaagatgagatgctgycggagaaagagcac 1219  
Db 1847 tctgtaatgcccagatgcatctgctgtttagatgagatcattgtgagaaagaaatc 1906  
Qy 1220 atgctgaactaattgcaaaaagagctcatattatctcctgtgatatgcaagatatta 1279  
Db 1907 atgtatgaaactatgaaagagagagcttacttcaactgttcaactgcaatgagaaagag 1966  
Qy 1280 aaaaagcgtgaca-acagatgagatcaatgacatctacttgaagaagaagacaaacca 1338  
Db 1967 gaaatgaaattgagttagaataatgccaactgtgtgaatccaaaagtgaaagtatgcttg 2026  
Qy 1339 ctccctcgcacatctgtgtaagagc-----atcaagctcagacttcatgcaagctg 1390  
Db 2027 aaatgctccaaaagatctcaggtccagtttaataaaagagatccaactcgcaagagta 2086  
Qy 1391 aggaatccaccacatcctaagaata-----agtc 1420  
Db 2087 tacatgcccacacagagccaaagacgaagcctgtgtacaaaagagagcttgaatgagatg 2146  
Qy 1421 ttcctgaagctcctctatataaattttaagttaaacgaagcctgaaagccttcttg 1480  
Db 2147 ttccttcagttcctctcctggagatctgaaagcgtgaactcaacgtgaagtgccctattt 2206  
Qy 1481 tctctgagacatggctctgttctcctaaatggaactgttcatccagatattcttcacatct 1540  
Db 2207 tgggtgataatcttgtccttataaagcgagcgctgcaacagacatttccaataatat 2266  
Qy 1541 ttgaaaaataataacatgctt---ggaataatgataaaacccatataagctagtg 1597  
Db 2267 ttccaagatataagggtctcttaccggagatgagagatccctgnaaacaagaagcagaata 2326  
Qy 1598 cagaattatcatcatgattctgcatcttgggtgatttggcttggcttgatcttca 1657  
Db 2327 gtaacatgttctctgtatgttctcagtccttggaattatttctttattacattttcc 2386  
Qy 1658 tgcagagatatttctacgagcagaggggaattttaacagatgagatlaagacatttgc 1717  
Db 2387 tccaaggtctacatttgcaagcgtgggagatccctcactaagggcttcgatatacag 2446  
Qy 1718 ccttcaaaagcatttataatcagatattgcctggttctgtagaagaaagaaacagacag 1777  
Db 2447 ttccaagatcatgtcagacagagatgcatgctgttctgtagaccccttaaaaacacactg 2506  
Qy 1778 gagcttgacacacatatattagccatagatatagacaaatccaagagcaacattcca 1837  
Db 2507 gagcatgacacacacagcctgtgccatgatatgctgctcaagttaaaagggtctatagttcca 2566  
Qy 1838 ggaatgctgtcttaacacaaaatgcaactaaactaagggacttccagttacattctctta 1897  
Db 2567 ggcctgctgtcatccacagaataatagcaaatcttggagacagcatattatataccttaa 2626  
Qy 1898 tatattgattgagatgacatctcctgattctggaatttgcctccagttacttgcgtgacag 1957  
Db 2627 tctatggttgcaattcaactacatttactcttagcaattgtaccatcatgtaacatagcag 2686  
Qy 1958 gaatgattgaaccgcgcgaatgactgattgtccacaagaagaaagcaaaactaagc 2017  
Db 2687 gagtgttggaaatgaataatgttctgtgacagacagcagaaagataaagaagctagaag 2746  
Qy 2018 atgctggaagaatagcaactgaaagccttggaataatagctactatgctgtaacaa 2077  
Db 2747 gagctggaagatgctacagaagccatcgaaaacttcggaactggttcttcttgatc 2806  
Qy 2078 gggaaaaagccttcgagcaaatgtatgaagatgcttcagactcaacacagaataacct 2137  
Db 2807 gggagcagaagttcgaatatactgacagagatttgcaagttccatacagaaactctc 2886  
Qy 2138 cgaagaaagcagatatttgaaagctgtltagatccagcagcttlatattttg 2197  
Db 2867 tgaagaaagcacaactctcgtgggtctcatcttctatcaccagcaatgtgtatttt 2926  
Qy 2198 cctatgcagcaggttctgatttgagccttaattcaagcttgagcaaatgacccag 2257  
Db 2927 cctatgctgctgttctcgtgttctgctactctgtgtgcaaaatgaaatctcagcttctc 2986  
Qy 2258 aggcagttatagtttttacttgcgaattgcatgtgagcctatgagcctcgaaaaagc 2317  
Db 2987 aggatgtcttctgtatctccacatctgcttctgtgagcagatgagtgagcaagttca 3046  
Qy 2318 tctgttggctcctgaaatattccaaagcaaatcggggctgcagatctgttgcctgt 2377  
Db 3047 gtcatctgctcctgcagactatgcaagagcaaatgatacagccacgtcatatgac 3106  
Qy 2378 tggaaaagaaccaaataatagacagccgagtcagaagaaggaagaaagccagacatgtg 2437  
Db 3107 ttgaaaaagccctctgtattgacagctacagccctcagcgctcctaaagccaaatcagttg 3166  
Qy 2438 aagggaatttagagtttcggaagatcttctctacatccatgctcgccagaatgtttca 2497  
Db 3167 aaggaatgtagacatttaagatgaggtgttccaactacatccactcgacacacatcccg 3226  
Qy 2498 tctcctgctgtatccctcagttatgagcgagaaagacagtagcatgtgtgggagca 2557  
Db 3227 tgcctcagggctgagcctcaggtgaaagagggccagacgctggccctcgttagtagca 3286  
Qy 2558 ggcgctgtgggaaaagcaactctgttcaactctgcagagacatttagaccctgtgcaag 2617  
Db 3287 gtgcctgtgggaaagagcaatttctcagctctcagagcgcttcatatgacccctgtg 3346  
Qy 2618 gacaagtgctgttgatgtgtgtagtgcataaagaattgaatgtaacagtggtccgtcc 2677  
Db 3347 gtccagtgctaatgtgagtgagcaagaagataaagcaccgtgaatgtccagtgtccgagcac 3406  
Qy 2678 aaatgacatcgttccctcaagagcctgtgctctccaactgaagcattgtgagacatcg 2737  
Db 3407 aacctgagatcgtgtcctcaagagcccatcctgttgaactgagatgtcgcgaagaattg 3466  
Qy 2738 cctatgttgacaacagccgtgtgtgtccatagatgagatcaaaagaagccgaatgtcag 2797  
Db 3467 cctatgagacaacagccgggtcgtgtatcacatgaaagatfatltcagagcagccaagag 3356  
Qy 2798 caaatcatctcttcttgaaggtctccttgagaaatacaaacacaaagtgtgacatga 2857  
Db 3527 ccaaatacacacattcatcgagacatcccttgagaaatacaaacacagaaatgagagaca 3586



OY	1220	atgtctgaactcaatgagcgaacaacgaggtctctatatctatctctgtgagtctcaagatctta	1279
Db	2259	atgtatgacccaacgaaagaaagaaagggatcttaacttcaactgttgcacaatgcagacgacg	2318
OY	1280	aaaaagctgtatgacaagatgagatcatgatcatctca-----ctga	1322
Db	2319	gaaatgaaattgttaattagaaaatgtcagcgtcgtatgaaatccaaaagtgaattgtccttgg	2378
OY	1323	aagaagaacccaactcacttcctctcgtcactctgtgaagagcatca-----	1366
Db	2379	aaatgcttccaatgatgtccaagatccagtcactcaataagaanaagatccaactcgttagagtg	2438
OY	1367	-----agtcagacttcattgtcacaagcgtcgtgaggaatccaaccaactcaagaagataatgc	1420
Db	2439	tccgtgtgattccacaagccccaagacagaagaagcttagtaccagaagagctctgtgataagta	2498
OY	1421	tctctgaaagctctctcattaaanaatttaaagttaaacaagcgtctaaagccttcttgg	1480
Db	2499	taacctcagttccctcttggagattatgaaagctcaaatttaactcgtaaagccttattctg	2558
OY	1481	tctcgggagacattggtcctctcgtctctcaatgaagacagcttcatccaglatcttccatcatc	1540
Db	2559	tgtgtgtgtattttgtgcccattatataaagtggggcgtcggaacagcatttgcataatat	2618
OY	1541	ttgcgaanaattataaccaaattgttggagaataatgatataaacaacataaag---cagtatg	1597
Db	2619	tttcaaaagattataggggttttcttacaagaatttgatgatctcttcaaaaagaaagacagata	2678
OY	1598	cagaatttatccatgatatactgcatacttctgtgtatgttctgtcttgcgttgcagttatcca	1657
Db	2679	gtaaactgttttcaactctgttctctcagcccttggaattatcttctttatcatcttcc	2738
OY	1658	tgcagcgattatttttaacgcagagacagaggggaaattttaacgatgaagattaagacacttgg	1717
Db	2739	tctcaggttccaacttcttgccaagaagcgtggagagctcctccacaagcgtctccgatacatbgt	2798
OY	1718	ctttcaagaagccatgttatatcatcagaatatgcccgtgttgatgataaagaaagcaagccag	1777
Db	2799	ttttccatccatcatgctcctaagacagagatgtgagtgtgttgaatcccttaaaaacaccctt	2858
OY	1778	gagcgttcacacacatatcttgacatgatagatagacaanaattcaagaaagacaagcttcca	1837
Db	2859	gagcattgactaccagagcctcgcgccaatgtagtctgtcctaagttaaaaggcctatagtctcca	2918
OY	1838	ggatctgcgttcttaaccacaanaatgycacactaacaatgggaatttcagttatcatcttccctta	1897
Db	2919	ggcttgcgttaattaccagagataatagcaaaactcttggcaggaataattatctccctta	2978
OY	1898	tatatgtatggtagagatgacattccctgattctcgtatgtctccagtaacttgcgttgacag	1957
Db	2979	tctatggtcttgacaactcaactacgttactctcttgaacatgtatgaccatcatctgcataatag	3038
OY	1958	gaatgattgaaacccgcgacaatgactgagatttttccaacaagaatagacaagaactaagc	2017
Db	3039	gagttgtgtgaaatgaaaaatgtgttcucgycgacaagcactgaaagaatagaagaagcaacgaag	3098
OY	2018	atgcctggaanaatagccaactgaaagccttggtagaataataagctcctatgtctaataaaca	2077
Db	3099	gtgcctgggaagatcgtcactcgaagcaatagaanaacttccgaacogtgtttctttgtactc	3158
OY	2078	gggaaaaagcctctcgagcaaatgtatagaaagatgctctcaagactcaacaacagaataactc	2137
Db	3159	aggagcagaagaattgtgaacacatatgtatgtcctaaggttttgagatgacatcacagaanaactct	3218
OY	2138	cgaaagaagaacagaaattatgtgaagcgtgtatgcatctgaagccatgcctttatataatttg	2197
Db	3219	tgaaggaagaacacacatcttgggaattacatttcccttccccaaggaatgatgataatttt	3278
OY	2198	cttaatgcagcaggggttctcgaatttttggagccattttaattcaagcctgcagacatgagaccag	2257
Db	3279	ccataagcgtgagatgttcccggttttggagccactactgtgtgacacataaaynmaagagccttg	3338
OY	2258	agggcacgttccataagatttttactcgaattgcatatgagacgtatggtccatcgtgaaacacgc	2317

Db	3339	aggatgtctcgttatgattcttcacgctgttcttgtagccatgycgctggggagatca	3398
Qy	2318	tcgtttgctcctcgtgaatatccaaagccaaatcggggtgctgacatctgttgcctgt	2377
Db	3399	gttattgtcctcgtactatgaccaaagccaaatatcagcagccacatcatatgta	3458
Qy	2378	tggaaaagaaccacaatatgacaagcccgatgcagaaagagaaaaagccagacatgt	2477
Db	3459	ttgaaaaaaccccttltgattgacagctacaagcaggaagccttaagtcgcgaacatgtg	3518
Qy	2438	aagggaatttagagtttcgcgaagctccttctctatccatgctgcgcagatgtttca	2497
Db	3519	aaggaataagtcacattgtgtgagaattggtatccaactatccaccgacccgacatccag	3578
Qy	2498	tcctccgtgcttalcctccatgatttgagcagagaaagacagtaagcatlttgtaggaaca	2557
Db	3579	tgtctcaaggagactgagccttgaggtgtagaagaaggccagacgctgtgctgtgtggaagca	3638
Qy	2558	gcgctgtggaagaagacattcgttcaactctgcagagacttaatacccgctgcaag	2617
Db	3639	gtgctgtggaagaagacagtggtccagctcctgtagaggtgttcacagcccttggcag	3698
Qy	2618	gacaaagtgcctttgtgtaggtgtgtagtcagaagaagattgaattgaacgtggtcctgtcc	2677
Db	3699	ggaaagtgcctgtgtgtagtggaagaagaataaagcgactgtgaatttcaagtgtctcggagac	3758
Qy	2678	aaatagcaatcgttctcctcaagagacgttgcttcaactgcagacattgctggaagaatcg	2737
Db	3759	acctggagatcgtgtgcccaaggagcccaactcgtttgtagctgcagacattgctggaacattg	3818
Qy	2738	cctatgtgtgacaagaacgccgtgtgtgtccattagatagaatcaagaagccgcaaatagag	2797
Db	3819	cctatgtgtgacaagaacgcccgtgtgtgtgtcacaagaagaagatcgttgagggcagaagaggag	3878
Qy	2798	caaatcatcattctttatttgaaggtctcccttgagaatacaacaacagaattgtgacctga	2857
Db	3879	ccaaatcatcattgcttcaatcgtgagctacgtcctataataatctgcacataagtagagaca	3938
Qy	2858	aagagcacaagcttcttcgctggccgagaaacaagaactagctatgtgcaaggtcctctcc	2917
Db	3939	aagagcattcagctctctgt	3998
Qy	2918	aaaaaacccaataatttatgttgtgtagtaggcacttcaagcctctcgatataagcaatgta	2977
Db	3999	gacacccctcatatttgccttcttggatbbaagccacgctcagctctgtgtataagaagtga	4058
Qy	2978	agggtgttcagcatgctcctgtatataaagccagggcgggaagacatgacctgagtgacatc	3037
Db	4059	agggtgttcacaagaagccctgtgacaagaagccagaaagagccgacctgtcatgtgattgtc	4118
Qy	3038	acaagctctctgcacatcacaagacagcatlttgaatgtgttctgcacaatgtgaaaagata	3097
Db	4119	accgctgtccacatccatccagatgacagacttaagtgtgttccgaatgacagatga	4178
Qy	3098	aggaaacaagaactcatcaagaagctcctgtgagaataatcgagacatataattttaagttagga	3157
Db	4179	aggagcatggaacgatcatagcagctgtgtgtgcacagaagaaggaacatctattttcaatgtga	4238
Qy	3158	atgcacagtcag 3169	
Db	4239	gtctcccaaggctg 4250	
<div> <div>RESULT</div> <div>7</div> <div>AAH57442</div> <div>AAH57442 standard; cDNA: 4349 BP.</div> <div>AAH57442;</div> <div>10-SEP-2001 (first entry)</div> <div>Human intestine cell specific cDNA sequence SEQ ID NO:282.</div> </div>			



Db 2320 ttccaagaatlaaggggtttttacaagaatltgatgatccttgaaacaaacagacagaata 2379  
 QY 1598 cagaaatlaattccatgatatctgcatttgggtgtatttcttctgttcagtatttca 1657  
 Db 2380 gtaactgttttaactacatttctctagcccttgaattatttctttattataacttttcc 2439  
 QY 1658 tgcagagatattttacgcgagagcaggggaaatttaacgatgagaattagaacttgg 1717  
 Db 2440 ttcaagggttccacatttggcaaaagctggagagatccctccacaagcgcctccgatactgg 2499  
 QY 1718 ccttcaagcagattatataccagatatctcctgggttggatgaaagaaacagacag 1777  
 Db 2500 ttctccgatccatgctcgaagagatgtgagttgttgaatgacctaaacacacactg 2559  
 QY 1778 gagctggaacaacataltgacagataatgacacaatttcaagagagacagagttcca 1837  
 Db 2560 gagcatgtaactacacagctcgcgaatgagtctctcaagttaaagggtccatagttcca 2619  
 QY 1838 ggaattggcgtcttaacaacaaatgcaactaacatgggaacttccagttatcaatttcttta 1897  
 Db 2620 ggcctgctgaattatccacagaatataagcaaatcttggacagagataattatattccttca 2679  
 QY 1898 tatatgagatgggagatgacattcctctgattctgagttatgtcccaagttacttggcgtgag 1957  
 Db 2680 tctatgtgttgcaactaacactgcttactcttaagcaattgtgacatccatcatggaatagcag 2739  
 QY 1958 gaatgtatgaaacccgacagcaatgactgatttgcacaacaaagataagacaagaacttaagc 2017  
 Db 2740 gagttgttgaataatgaaatgctgtctggacagagcactgaaagataagaagaacttagaag 2799  
 QY 2018 atgctggaagatagcaactgaagccttggagaaatatacgtactatgtgtcattaacaa 2077  
 Db 2800 gtgctgggaagatcgcactactgaagcaatagaaacttccgacacgtgttcttcttgatc 2859  
 QY 2078 gggaaaaagccttcgagcaaatgtatgaagagatgcttccacaacacagaaataact 2137  
 Db 2860 aggaagcagaagtttgaacataatgtatgctcagagtttgcagatccataacagaaactct 2919  
 QY 2138 cgaagaagcagacagatatttgaagcgtgtatgcatltagcactgaagcacttataatttgg 2197  
 Db 2920 tgaagaaagcacacactcttgaattataatttcttccaccaggaatgtgtatttt 2979  
 QY 2198 cctatgcaagcagggttctgatttggacactttaaactcaagctggacgaatgacccag 2257  
 Db 2980 cctatgctgagttcttcgcgttcttggagcctacttgggtgacataaactcatgaccttgg 3039  
 QY 2258 agggcagtttcatagtttttctctgcaattgcatatgagcgtatggccatcggaaaaaagc 2317  
 Db 3040 aggaatgtcgtttagtatttccagctgttcttcttggcgtgacatggcgtgggagcaatca 3099  
 QY 2318 tctgttggcctccgaatatattccaaagcaaatcggggcgtgcgcatctgttggccttgt 2377  
 Db 3100 gttcatttgcctccgactatgccaagcacaataatcagagccacatcatcatgataca 3159  
 QY 2378 tggaaagaagaaccaaatalagacagcgcagtlcaaagaagggaagaaacagacacatgtg 2437  
 Db 3160 ttgaaaaaaccccttggattgacagctacagacagcaggaagccttaatgpcgaacacattgg 3219  
 QY 2438 aaggaatttgaagtttctgaaagatccttcttctatccatgttgcagagatgttttca 2497  
 Db 3220 aaggaatgtacacatttgttgaagttgtatcaactlcccccagaccgagacatccag 3279  
 QY 2498 tccctcgtggttattccctccagtatgtgagcgaagaaagacagtaacatttgtgggagca 2557  
 Db 3280 tgcctcagggagactgacctggagagtgaaagaaggccacaagcgtgctctctgtgggagca 3339  
 QY 2558 gcgcgtgtgggaaagacacttctgttcaactctgcagagacttataaccccggtgcaag 2617  
 Db 3340 gtgcgtgtgggaaagacagatgtccagctcctgtgagcgtgttcaagcacccttggag 3399  
 QY 2618 gacaagtgctgttcttgaatgtgtgagtgcaaaagaattgaattgaagtgctcgttccc 2677  
 Db 3400 ggaagtgctgctgtatgtgcaaaagaataaagcagctgaattctagtgctccgagac 3459

QY 2678 aaatagaatcgttcttccaaagcctgtgctcttcaactgacgacattgtggaacatcg 2737  
 Db 3460 acctgggcacgtgtccacaggagccatctctgttactgacgacattgtggaacatcg 3519  
 QY 2738 cctatgtgacaaagccgtgtgtgacattagatgacataaagaagccgacaatgacg 2797  
 Db 3520 cctatgagacaacagccgggtgtgtlcaaggaagaagatcgtgagggcagcaagagag 3579  
 QY 2798 caatatccattctttatgtgaaggtcccttgagaataacaacaaagtttgagctga 2857  
 Db 3580 ccaacatatactgtcttcatcagtagtcttactgtcctaataataagcactaaagttagagaca 3639  
 QY 2858 aaggaagcagccttcttgcggcgccagaaacaaagactagctatttgaaggcctcttcc 2917  
 Db 3640 aaggaatccagctctctgtggtggccagaaacaaacagcattgccaagtcgtgacctgtta 3699  
 QY 2918 aaaaacccaatttatttgttgatgagccacttccagcctcgtatataatgacagtga 2977  
 Db 3700 gacagcctcatatttcttcttggatgaagccaagtcagctcgtgatacagaagtgaaa 3759  
 QY 2978 agtgtgtcagcaltgccttgataaagccagagcgggaagacatgctctagtgtaactc 3037  
 Db 3760 aggtgtccagaagcccttgacaaagccagaagagccgacactgcatgtgtatgtctc 3819  
 QY 3038 acaggtcctctgcaattcagaacagcagatttgaatgagtggtctgcacaatgggaagtaa 3097  
 Db 3820 accgcctgcacacatccagaaatgacacttaatagtggtgttccagaaatggcagagta 3879  
 QY 3098 aggaacagagacactcagcagagctcctgagaatctgagacatatatttaagttaga 3157  
 Db 3880 aggaacatggcagcagcagcagcagctgtgtgacagaaagcagcatctattttcaatgtca 3939  
 QY 3158 atgcacagtcag 3169  
 Db 3940 gtgcacaggtcg 3951

RESULT 8  
 AA072872 ID AA072872 standard; cDNA; 4646 BP.  
 XX  
 AC AA072872:  
 DT 21-JUN-1995 (first entry)  
 XX  
 DE Human multidrug resistance gene (MRD-1) encoding a p-glycoprotein.  
 XX  
 KW Human; tyrosinase; compound; hair; follicle; liposome; lipophilic;  
 KW lipophobic; stratum corneum; melanin; hair dye; glycoprotein; alopecia;  
 KW chemotherapy; transforming growth factor; growth stimulant; aromatase;  
 KW cyclosporin A; macromolecule; polymer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 425..4267  
 FT /\*tag= a  
 FT /product= p-glycoprotein  
 PN W09422468-A.  
 PD 13-OCT-1994.  
 XX  
 PF 01-APR-1994; 94WO-US03634.  
 XX  
 PR 02-APR-1993; 93US-0041553.  
 PR 13-JAN-1994; 94US-0181471.  
 XX  
 PA (ANTI-) ANTICANCER INC.  
 XX  
 PI Li L, Lishko VK.  
 XX

DR WP1: 1994-332816/41.  
DR P-PSDB: AA072872.  
XX Liposomes for delivering protein, nucleic acid etc. to hair  
PT follicles - e.g. to restore hair colour, prevent hair loss during  
PT chemotherapy, stimulate hair growth etc.  
XX  
PS Claim 8; Page 70-77; 100pp: English.  
XX  
XX The nucleotide sequence of the human multidrug resistance (MDR-1) gene  
CC which encodes a P-glycoprotein. This is an example of a compound which  
CC can be delivered to hair follicles via a novel liposome composition.  
CC The liposomes are comprised of a lipophilic or lipophobic compound which  
CC will selectively target the hair follicle (via the stratum corneum)  
CC without damaging or unwanted effects on cells outside the follicle.  
CC Compound e.g. P-glycoprotein, can be delivered to the hair follicle for  
CC treatment of chemotherapy-induced alopecia. Other compounds targeted at  
CC hair follicles can include: tyrosinase (or the DNA encoding it -  
CC AA072871), melanin or hair dyes (to restore hair colour or condition);  
CC human transforming growth factor-alpha (AA072873) (for reversal of wavy  
CC hair); hair growth stimulants (e.g. cyclosporin A or arimataze) or  
CC antisense sequences. The method allows compounds (e.g. macromolecules  
CC or polymers), which would not normally reach the hair follicles, to be  
CC delivered to these target areas.  
XX  
SQ Sequence 4646 BP: 1371 A; 893 C; 1128 G; 1254 T; 0 other;  
  
Query Match 35.4%; Score 1124; DB 15; Length 4646;  
Best Local Similarity 62.1%; Pred. No. 6.1e-294;  
Matches 1869; Conservative 0; Mismatches 1095; Indels 48; Gaps 4;  
  
OY 200 ttgtatgatacacagatctcaagaatgcaaggatttggcattaaaggactatag 259  
DB 1245 ttgaaagttaacacaaaatttagaagaagctaaagaattggatagaagaagctatta 1304  
OY 260 cttaaaaggtgctctgtgctgtgtgtactcttcttaagatggaacctatgactgtctt 319  
DB 1305 cagccaatattctatagtgctgtcttccgtcgatctatgacattatgtctcgacct 1364  
OY 320 ttggataggaacctccttgatcttctaaatggaacctggaatataccatcggaactgttc 379  
DB 1365 tctgtatggagcacacctgtgctctcctcagggga-----atattctatggacaagtaac 1418  
OY 380 ttgtctgttcttctatgtatcatcagatgatttgcattggagcagcagctccctact 439  
DB 1419 tcaactgtatcttcttctgtatattgggctttagtgttgacagacatctccaaga 1478  
OY 440 ttgaaccttcgaatagcccgagagctgctcttcatatttccaggttatgttaaga 499  
DB 1479 ttgaagcatctgcaaaatgcaaaagagcagcttaagaatctccaagaatctgaataa 1538  
OY 500 aaccagatatagaatactttccacagcttgaataaaacctgaatccatagagaactg 559  
DB 1539 agccaagatctgacagctattcgaagagtgagcacaacacagataatattaaaggaaatt 1598  
OY 560 ttgaatttaaaaatgttcttcaattatcacaagaacacatctcaagaattctgaag 619  
DB 1599 ttgaattcaaaaatgttcaattatcacaatctccgaaaagaattgaagcttctgaag 1658  
OY 620 gtctaaactcagaattgaatctggaagacagctgccttggctggcttccaaatggcagtg 679  
DB 1659 gcttgaacctgaaggtgcaagatggcgagagcggtggccttggttggaaaactgtgctgtg 1718  
OY 680 ggaagatagcgttagctcagacttctgcagaggtatataatgctcggatgtatgctttatca 739  
DB 1719 ggaagagcaacaacagctcagctgacgagctctatgaccccaagagggatgtgtca 1778  
OY 740 tgggtgtagagagatgacatcagaacttaaatgtgcggcatcatatcagacatattggag 799  
DB 1779 gtgtgtgtagagagatattttagaccataaataatgtaagtttctactaggaataatctgtgt 1838  
OY 800 tggtagtcaagagccttctgttcgggacacacatcagtaacatatacgaattatgac 859

DB 1839 tggtagtcaaggaacctgtatgttctgcccacacagatagctgaaacattcgtctatggcc 1898  
OY 860 gagaagatgtgactgatgaagagatgagagcagcaagggaaagcaaatgctatgatt 919  
DB 1899 gtgaataatgtccacttgatgagatttggaagaagctgtcaaggaagccaatgctatgact 1958  
OY 920 ttatcatgagtttccctataataattatcatctgttaggggaaaaaagagctcaatga 979  
DB 1959 ttatcatgaaactgtccctataaatttgacaacctgtgttgagagagagggccagcttga 2018  
OY 980 gtgagggagcaaaaagagagatccgcaatgtgctgctgacttagtggaaacccaagatttc 1039  
DB 2019 gtgtgtgagcaagaagagagatcgcatctgacgcgtgacctgtgttcgcaacccaagatcc 2078  
OY 1040 tgattttgagttagctacgtctgctccctggaattcagaagcaagtcagctgttcaagctg 1099  
DB 2079 tctgtgtgagttaggcacagctcagctcgttgacacagaagaagcgaagctgttcaagctg 2138  
OY 1100 cactgtgagagcgagcaagagctcgactacaatcggtgttagcaacacgactttacta 1159  
DB 2139 cctgtgataaagcagaaaaggttcgaccacacatgtgatacgtatcgttgcctacaag 2198  
OY 1160 ttcgagatgcaagatttgattgtgacctaaagatggaatgtgtgcggagaaagagacac 1219  
DB 2199 ttctgataatgtacgtctacgtctgtgttctcogatagtgagtgctgttggagaaggaatc 2258  
OY 1220 atgtcgaactaaatggcaaaaagagctctatattatcactgtgtatgacagatatha 1279  
DB 2259 atgtgacactccttgaagaaggaagacatttcaacttccaaatgacagacagcag 2318  
OY 1280 aaaaagctgatacagaatgagatggatgaatgacatatttca-----ctga 1322  
DB 2319 gaataagatgtgaattagaataatgcaagctgataatccaaatgaaatltgactgtgag 2378  
OY 1323 aagaagac 1366  
DB 2379 aaatgtcttcaaatgattccaagaatccagctcgaatgaagaagaaagacacacacacacac 2438  
OY 1367 -----agtcagacttcatgcaaaagctgagaaatccacccaatcctaaagagataagtc 1420  
DB 2439 tccgtgatacacaagcccaagacagaagacttagtaccacaagaagctctgtatgtaaga 2498  
OY 1421 ttcctgaagctctctctataaaaaatttaaaatttaacaaagcctgaatgcttctgtg 1480  
DB 2499 taacctcagtttcccttcttgagagattatgaactaaatttaactgaatgtgcttatttg 2558  
OY 1481 tcttgaggaactgtgcttctgttcttaaatggaactgttcatcagaatttccatcatc 1540  
DB 2559 ttgtgtgtgtatttctgtcattataaattgagagctgtgaacacagcatttgcataaact 2618  
OY 1541 ttgcaaaaataaaacacatgttggaaaataatgataaaaacacatthaaag---catgag 1597  
DB 2619 ttcaagaatataaggggttttacaagaattgatactcctgaacaaacagacagaata 2678  
OY 1598 cagaattatcattcattatcattcgtcatttgggtgtatttggcttggcatttttca 1657  
DB 2679 gtaactgttctcactatgttctttagcccttggaattatttcttattatcaatttctc 2738  
OY 1658 tgcaggaattatttcaagcagagagaggaattttaaagctatgatatgataacacactg 1717  
DB 2739 ttcagagtttccacttggcagaagctgagagatcctcaacaaagggccgcatatacag 2798  
OY 1718 ccttaaaagcactgtatatacagatvatgtgctgtgttgaatgaaaaagaaagacagacag 1777  
DB 2799 ttcttcgataccttccagacagatgtgtgtgttttagatgaccctaaacaaacacactg 2858  
OY 1778 gaggcttgaacaacataatagccatagatatagcaacaatttcaagagagaaagcttcca 1837  
DB 2859 gaggatctgactacagagctcgcccaatgagctgtctcaagtttaaaggggtactaagttcca 2918  
OY 1838 ggaatggcgtcttaacacaaaatgcaactcaacatgggaattcagttacattcttcctta 1897

Db 2919 ggtctgtctaataaccagaataatagaacaaactctgtggacaggaataataatatacttca 2978  
 QY 1898 tataatgagtgagatgagatgagatcttctgagatcttctgagatctgagcagtgagacag 1957  
 Db 2979 tctatgtgttgcaactaacaactgttactcttaagaatctgtaccatctatctgcaatagcag 3038  
 QY 1958 gaatgatgaaacccgcagcaatgactgtattgtccaaacaaagaataagaactaaagc 2017  
 Db 3039 gactgtgtaaatgaaatgactgtctgacagaagcaactgaagaagaactaagag 3098  
 QY 2018 atgtctgaaagatgacactgaaagcttctggagaataatagctactatgctgactaa 2077  
 Db 3099 gtgctgggaagatgcactgttaaagcaataagaaactccgaaccgtgttcttctgac 3158  
 QY 2078 gggaagaagccttcgagcaaatgcatgaaagatgagcttcaactcaacacagaataact 2137  
 Db 3159 aggaagcaagaatgtgaacatctgtatgtctcagagtttgcaagtacacacagaactct 3218  
 QY 2138 cgaagaagacagagattatggaaagctgttatgcatcgaacagcacttataatttg 2197  
 Db 3219 tgaagaaagcacacactcttgaaattacatttctcaccaggaagcaatgtatgtt 3278  
 QY 2198 cctatgcagcaggttctcgatttgagccttataatcgaagcctggagcaatgacccag 2257  
 Db 3279 cctatgtgagatgtcttcggttgcagctactctgtgacataaactcagacttg 3338  
 QY 2258 aggcacatgcatagtttctactgcaattgcatatgtgagacatgccaacgcgaacacgc 2317  
 Db 3339 aggatgtctctgtatgtatcttcaagctgtgtctgtgtgcacatgvcgtgggacagca 3398  
 QY 2318 tctgttgcctcctgaatattccaaagcaaatcggggctgcagctgttgccttctgt 2377  
 Db 3399 gtcatctgtcctcactatgcgaagcaaatatcagcagccacatcatcatgac 3458  
 QY 2378 tggaaagaacacacaaatataacagccgcagatcaagaagggaaaaagccacacatgtg 2437  
 Db 3459 ttgaaaaaaccccttgatgtacagctacagacgaagcgaagcctaattgcgcgaacatgtg 3518  
 QY 2438 aagggaatttgaagttctgaaagatctcttctctacatgctgcgaacatgtttca 2497  
 Db 3519 aaggaaatgtcacatctgtgtaagtgtattcaactaccacccgagacgaacatccac 3578  
 QY 2498 tctcctgtgctatccctcagatlttgacgagaaagacagatgcatgtgtgggagca 2557  
 Db 3579 tgtctcagggactgagcctgagagtgaaagggccagacgtgctgtgtggagca 3638  
 QY 2558 gcgcgtgtgggaaaaagcactctgttcaactctgcagagagacttatgacccgtgacag 2617  
 Db 3639 gtgctgtgggaaagacagcagtgccagctcctggagcggcttcaagaccccttgagcag 3698  
 QY 2618 gacaaagtgtctgtgatgtgtgatgcaaaagaattgaaatgataagtgagctgcctcc 2677  
 Db 3699 ggaagtgctgtgtgatgcaaaagaataaacgcagatgctcagtgctgcagagac 3758  
 QY 2678 aaatagcaatcgtctcccaagaagcctgtgcttcaactgcagcagatgtctgagaacatcg 2737  
 Db 3759 accgtgcatcgtgtcccaagagccacccctgttgcagatgtctgagaacatcg 3818  
 QY 2738 cctatgtgacaacagccgtgtgccaatagatlgaaataaagaagccgcaaatgag 2797  
 Db 3819 cctatgtgacaacagccggtgtgtgtcaagagaaatctgtgagggcgaagaagag 3878  
 QY 2798 caaatatcatcttctttaaagtgctcctcctgagaataacaacacaaattgagacga 2857  
 Db 3879 ccaacatcatcagccttctatcagatcactgactataataatagcaactaaatagagaca 3938  
 QY 2858 aaggaacagacgtctcgtggcgccagaaacaaagactatgtgcaagggccttctcc 2917  
 Db 3939 aaggaactcagctcgtgtgtgccaagaacacagcatgtcatgctgtgccttgtta 3998  
 QY 2918 aaaaacccaattttatgttgatgagggcacttccgctcctgataatagacagtga 2977  
 Db 3999 gacagcccatatttcttcttgatgaggaacacagctcagcttcatagacaagaattgaa 4058

QY 2978 agtgtgtcaacatgccttgcataaagccagagcggaagacatgctagtgacac 3037  
 Db 4059 aggtgtccaagaagcccttgacaaagcagaagagccgacactgctatgtgtctc 4118  
 QY 3038 acaggtctctgcaattcagaacgagatttgatagtggtctgcacaaatgaaatga 3097  
 Db 4119 accgctgtccaccatcagatgcagactaaatagtggttccaaatgycagatga 4178  
 QY 3098 aggaacaaggaactcatcaagagctcctgagaagaatcgacatatatttaagtga 3157  
 Db 4179 aggaagatggacagcagatcagcagctgtgcacagaagaagcattctatttcaatgtca 4238  
 QY 3158 atgcacagtcag 3169  
 Db 4239 gtgtccaggtcg 4250  
 RESULT 9  
 AA294738  
 ID AA294738 standard; cDNA; 4646 BP.  
 XX  
 AC AA294738;  
 DT 01-AUG-2000 (first entry)  
 XX  
 DE Human ATP binding cassette ABCB1 (MDR1) cDNA.  
 KW ABCB1, ATP binding cassette; human; cholesterol; lipid disorder;  
 KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;  
 KW lupus erythematosus; diagnosis; gene therapy; MDR1;  
 KW multidrug resistance; chromosome 7q21; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 EN M0200018912-A2.  
 XX  
 PD 06-APR-2000.  
 XX  
 PF 21-SEP-1999; 99WO-EP06991.  
 XX  
 PR 25-SEP-1998; 98US-0101706.  
 PA  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Schmitz G, Klucken J;  
 DR WPI: 2000-293151/25.  
 XX  
 PT Adenosine triphosphate binding proteins useful for identifying agents  
 PT for treating atherosclerosis and other inflammatory disorders -  
 PS Claim 9; Page 110-112; 154pp; English.  
 XX  
 CC The present sequence is that of human ATP binding cassette  
 CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a  
 CC differential display method in which monocytes from peripheral  
 CC blood were subjected to macrophage differentiation and cholesterol  
 CC loading with acetylated low density lipoproteins and subsequent  
 CC degrading with high density lipoprotein (HDL) to identify  
 CC cholesterol sensitive genes. The gene maps to chromosome 7q21  
 CC and is also termed MDR1 (multidrug resistance). The invention  
 CC provides cholesterol-sensitive ABC genes (see AA294734-63). These  
 CC genes, and polypeptides encoded by them, can be used for diagnostic  
 CC and therapeutic applications, and for biochemical or cell-based  
 CC assays to screen for pharmacologically active modulator compounds  
 CC useful for the treatment of lipid disorders, atherosclerosis or  
 CC other inflammatory diseases such as psoriasis and lupus  
 CC erythematosus.  
 XX  
 SQ Sequence 4646 BP: 1371 A; 892 C; 1129 G; 1254 T; 0 other;









Dh 1599 tggaaatcagaatgttcaacttcagttacccatctccgaagaagttaagatcttgaag 1658  
Qy 620 gtctgaatctcagaattgaatcttgagagacagtcgccttggttcgtccaatgacgtg 679  
Dh 1659 gccctgaacctggaaggtgcagagtggtgcagacggtgccttggttcggaacaatggtcgtg 1718  
Qy 680 ggaagaagtcaggttaagtcacagcttcgcagaggttataatgatccggatgtgtttatca 739  
Dh 1719 ggaagaagcacaacagatccacagctgcagaggtctctatgacccccaagaggtgagttca 1778  
Qy 740 tgggtgagagaatgacatcagaggtttaaattgtgcggcatatctcgagacatattgag 799  
Dh 1779 gtgttgatgagacaggtattgaaccataaattgtaagtttctaagggaatcatgtgtg 1838  
Qy 800 tgggtatgcacaagacctgtttgttcgcggagccacaatcagttacaatcaatgattgac 859  
Dh 1839 tgggtatgcacaagacctgtttgttcgcacaacacagatagcttgtaaaacatctgccttg 1898  
Qy 860 gagatgagtcagctgatatgaagatggaagagcagcagaagggaaacaaattgctatgatt 919  
Dh 1899 gtgaaatgtccacatgataatgagattggaagagctgtcagaagagccaaatgcctatgact 1958  
Qy 920 ttatcatgagtttccataataatttaacatctgttgagggaagaaagagtcacaatga 979  
Dh 1959 ttatcatgaaactgcctcatataatttgacaccccttggttgagagagagggccagttga 2018  
Qy 960 gtggagggcagaagaagagagatgcgaattgctgcgttccttagttcggaaaccccaagattc 1039  
Dh 2019 gtgggtggcagaagagagagatgcgcacatgcgcgtgccttggttcgcacaccccaagattc 2078  
Qy 1040 tgaatttgatgaagctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgc 1099  
Dh 2079 tccctgcgtgatgaggtccagcgttcagctgcctgcctgcctgcctgcctgcctgcctgcctgc 2138  
Qy 1100 cactgagagaagcgcagcaaaaggttcgagactacaatcgtgtgtagcaaacgacttctacta 1159  
Dh 2139 ctctctgataagccagaaaggttcgcgacacatctgtgtagctatcactgttgccttaacg 2198  
Qy 1160 ttcggaagtcgagatttgaattgtgaccccaagatggaatgctgcgcgggaagaagacac 1219  
Dh 2199 ttcgtaaatgtgacgtctcgtctggtttcogatgagtgatctatctgtggaagaagaaatc 2258  
Qy 1220 atgcggaactaattgcaaaaacgagctcactatctactatcactatggttgcacagagatata 1279  
Dh 2259 atgaggaactcagaaagagaaagacattacttcaactgtgcacaatgcagacagacg 2318  
Qy 1280 aaaaagctgagacagatgtagatgacatgacatctca-----ctga 1322  
Dh 2319 gaaatgaaattgaaattgaaataatgcagctgagatccaaaagtgaattgtgccttgg 2378  
Qy 1323 aagaaagacaaactcaactctcctctgcactcgttgaaagagcatca----- 1366  
Dh 2379 aaatgtctcaaatgattcaagaatccagctcactataagaanaagatcaactcgttaagatg 2438  
Qy 1367 -----agtcgaactctattgacaaggtcgtgaggaatcccaactcaagaagatgaagtc 1420  
Dh 2439 tccgtgagatcacaagcccaagacagaagacttaagraccaagaaggtcctgcgttgaaagta 2498  
Qy 1421 ttcctgaagctctctatctaaatatttaagaattaaacaaagctgtaatgaccttctgtg 1480  
Dh 2499 taaccacagtttcccttcttgagagattatgaagctaaatttaactgaatgaccttatttg 2558  
Qy 1481 ttcctgggagatgtgcctctgttcttaaatgtgaactggttcacatcagatatttccacatc 1540  
Dh 2559 tctgtgtgtgtattgttgccattataaagtgaagcgtgcgaacacagacatttgcaataat 2618  
Qy 1541 ttgcataaataataacacgtgttggaataatgaataaagccacatnaag---catgagt 1597  
Dh 2619 ttccaagaattcaggggttttacaagaatgtagatccctgaacaaacagacagaata 2678  
Qy 1588 cagaaatctatcagatgatatcgtcattttgggtgttattgtgcttgcagtattca 1657  
Dh 2679 gtaactgtttcaactattgtttctagcccttggaattattcttattatatttcc 2738

Qy 1658 tgcagggattattttaacgacagacagggaaattttaacgattgaatgaagacacttgg 1717  
Dh 2739 ttcaagggttccacatttgcgaagcgttgagagatccccaagggccgtgatacag 2798  
Qy 1718 ccttaaaagccatgttataatccagatattgcctgttttatatgaagaaagaaacagacag 1777  
Dh 2799 ttcttcagatcactgctcagacagaggttgagttgttttgatgacccctaaacacacacg 2858  
Qy 1778 gaggtctgaacaatataatgacatagatalagcaacaatccaaggaacaggtttcca 1837  
Dh 2859 gagcatctgactaccagcgtgcgaatgactgcctcaagttaaagggtataggttcca 2918  
Qy 1838 ggaattggcgtcttaacacaaatgcaactaaatggaattggaattcagttatcatcttcta 1897  
Dh 2919 ggccttgtaattatccccaatataagcaaatcttgggaaggaataatataatcccttca 2978  
Qy 1898 tatatgagatggagatgacatccctgattctgagatttctccagatctgcgtgacag 1957  
Dh 2979 tctatggttggaactaaactaaacgttactctcttagcaattgttaccatcatctgaatagcag 3038  
Qy 1958 gaatgattgaacccgcaagcatgactggaatttgcacaacaagaatagaagaactaaagc 2017  
Dh 3039 gagttcgtgaatgaaatgaaatgttctgcgacaagcactggaagaatagaagaactagaag 3098  
Qy 2018 atgcgtggaagatagcaactggaacttggagaatataagttctatagtctataacaa 2077  
Dh 3099 gtgcgtggagaagctgcctcaatgaaacatagaaaacttcggaacccgttgtttcttggaac 3158  
Qy 2078 gggaaaagaccttcagacaaatgatatgaagagatcttcaagctcaacacagaataact 2137  
Dh 3159 aggaagcagaagtttgaacatagttatgctcagaagtttgaaggttacatcacaaactct 3218  
Qy 2138 cgaagaagcacagattatgtgaagctgttatgcatccaagcagccttataatttgg 2197  
Dh 3219 tgaagaaacacacactcttgaattatacttctccctcaacagcgaatgatatttt 3278  
Qy 2198 cctatgcagcaggttttcgatttggagcctatcttaattcaagcttgagcaagaccacg 2257  
Dh 3279 cctatgctgaaatgttccgttctgagcctacttgttgcacataaactcagacacttgg 3338  
Qy 2258 aggcacatgcatagtttttactgcgaatgcatatggaacttgcgcacatgcggaagaaagc 2317  
Dh 3339 aggaatgcttctgaattttcaagctgtgtcttgggtccatgcggtggggcaagttca 3398  
Qy 2318 tgcgttgcctcgaatatcccaagccaaatcgggggtgcgcatctgttgcctgt 2377  
Dh 3399 gtccaatttgcctcgtactatgcgaagccaaatatacagcagcccaatcatcatgatca 3458  
Qy 2378 tggaaagaacccaatatagacagccgagtcagaagaagggaaagccagacacatgtg 2437  
Dh 3459 ttgaaanaaccccttctgattgacatagacagcaggaagccttaatgcccgaacaattgg 3518  
Qy 2438 aaggaatttagatttgcgaagctcctcttcttcaatcagctgcgcgaagatgttttca 2497  
Dh 3519 aaggaatgtccacttgggtgaagttgtatccaactatccaccggaacccgacalocceg 3578  
Qy 2498 tccctcgttgcttactccctcaglattgagcgaagaagacagtlagatatttggggagca 2557  
Dh 3579 tgcctcaaggactgcgcctgcgaggtgaagaagggccaagcgtctgcgtgcggagca 3638  
Qy 2558 gcgcgtctggggaagacacttctgttcaactctcgcagagacttataagccccggagcaag 2617  
Dh 3639 gtgcgtctggggaagacagatggtccatccctctggaagcgtgttcttaagcccccttggagc 3698  
Qy 2618 gacaagctgcttgaatggctgcgaatgcaaaagaaattgaaatgataagctgccttcc 2677  
Dh 3699 ggaagatgctgtcttgatgagcaagaataaagcagatgaaatgttcaagtgcccgagac 3758  
Qy 2678 aatagcaatcgttcccaagacgtgtgcttccaactgcagacatgtctgagacatcg 2737  
Dh 3759 acctggacatcgttccacagagccaatcctgtttgacgtcagcatgtctgagagacatg 3818

QY 2738 cctatggtgacacacagcgtggtggtcatttagatgataagaagccgcaatgacag 2797  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 3819 cctatgagacacacagcgtggtggtcatttagatgataagaagccgcaatgacag 3878  
2798 caaatcattcatttatttgaaggtctccctgagaataacacacaggttgagactg 2857  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 3879 ccaacatcattcatttgaaggtctccctgagaataacacacaggttgagactg 3938  
2858 aagagacacagcttcttggtggtggtggtggtggtggtggtggtggtggtggt 2917  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 3939 aagagacacagcttcttggtggtggtggtggtggtggtggtggtggtggtggt 3998  
QY 2918 aaaaaccacaaatttatttgaaggtctccctgagaataacacacaggttgagactg 2977  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 3999 gacagcctcatttcttcttcttcttcttcttcttcttcttcttcttcttcttct 4058  
QY 2978 aggtgtgtcagcagcttcttcttcttcttcttcttcttcttcttcttcttcttct 3037  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 4059 aggtgtgtcagcagcttcttcttcttcttcttcttcttcttcttcttcttcttct 4118  
3038 acagagctctcagcagcttcttcttcttcttcttcttcttcttcttcttcttcttct 3097  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 4119 acagagctctcagcagcttcttcttcttcttcttcttcttcttcttcttcttcttct 4178  
QY 3098 aggaacacaggaacacacagcagcttcttcttcttcttcttcttcttcttcttcttct 3157  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 4179 aggaacacaggaacacacagcagcttcttcttcttcttcttcttcttcttcttcttct 4238  
QY 3158 atgacagctcag 3169  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 4239 gtgtccagcgtg 4250

RESULT 11  
AAT13394  
ID AAT13394 standard; DNA: 6505 BP.  
AC AAT13394;  
XX  
DT 24-JUN-1996 (first entry)  
XX  
DE Hybrid vector pSF-MDR.  
XX  
KW Hybrid; vector; gene transfer; gene therapy; haematopoietic stem cell;  
KW retrovirus; murine embryonic stem cell virus; MESV;  
KW Moloney murine sarcoma virus; (MOMuSV);  
KW Friend murine leukaemia virus; F-MuLV; ds.  
XX  
OS Synthetic.  
XX  
PN DEL19503952-A1.  
XX  
PD 14-MAR-1996.  
XX  
PE 07-FEB-1995; 95DE-1003952.  
XX  
PR 08-SEP-1994; 94DE-1431973.  
XX  
PA (BOF ) BOEHRINGER MANNHEIM GMBH.  
XX  
PI Baum C, Osterlag W, Stocking-harbers C, Stockingharbers C;  
XX  
DR WPI; 1996-152306/16.  
XX  
PT Hybrid retroviral vectors - for gene transfer into haematopoietic  
XX stem cells  
XX  
PS Disclosure: Page 25-29; 42pp; German.  
XX  
CC New hybrid vectors comprise (1) a leader region including the U5  
CC region and tRNA primer binding site of murine embryonic stem cell  
CC virus (MESV) or Moloney murine sarcoma virus (MOMuSV), and (2) a 3'-  
CC LTR including the U3 and R regions of a Friend murine leukaemia

CC virus (F-MuLV). The vectors are useful for ex-vivo or in-vivo gene  
CC therapy. High levels of gene transfer can be achieved in  
CC haematopoietic stem cells and their myeloid (non-lymphatic) progeny.  
CC pSF1, pSF2, pSF3 and pMML (sequences given in AAT13390-T13393) are  
CC examples of such vectors.  
CC Vector pSF-MDR (sequence given in AAT13394) is based on the  
CC MESV vector R224.  
XX  
SQ Sequence 6505 BP; 1719 A; 1506 C; 1606 G; 1674 T; 0 other;

Query Match 35.4%; Score 1124; DB 17; Length 6505;  
Best Local Similarity 62.1%; Pred. No. 7,3e-294;  
Matches 1869; Conservative 0; Mismatches 1095; Indels 48; Gaps 4;

QY 200 ttgtaaggtatcacagaaatctcacaagatgcaagatttgcgataaaaggactatag 259  
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2637 ttgaaggtacacaaanaatttgaagaagctaaagaatttggataaagaagactatta 2696  
QY 260 ctccaagagctctcttggtggtggtggtggtggtggtggtggtggtggtggtggt 319  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2697 cagccaatttctataggtggtggtggtggtggtggtggtggtggtggtggtggt 2756  
QY 320 ttggtatgaaacctcttgattcttcaatgagaacctgataaccatcgagactg 379  
||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2757 tctggtatggagaccaccttggtcctccagggg-----atatctatggagcaagtac 2810  
QY 380 ttgctgttcttctttagtgaatccatagcagttatgcaatttggagcagcgtccctact 439  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2811 tcaactgtattcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2870  
QY 440 ttgaaccttcgcaatgcccggagagctgcttccatatttcccaaggttatgataaga 499  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2871 ttgaagcatttgcacaagagagagcagctatgaaaccttcaagataatgataata 2930  
QY 500 aaccagatagataaactttccacaagctgataataaacttgaatccatagaaagac 559  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2931 agccaagattatgacagctatgcaagagtggtgacacaacagataatlaagagaaatt 2990  
QY 560 ttgaatttaaaatgcttcttcaattatccataaagaccatcataagattctgaaag 619  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2991 tgaattcagaagaatgcttcaacttcaacttcaacttcaacttcaacttcaacttcaact 3050  
QY 620 gtctgaattcagaattgaatgaatgagagagcagctggtggtggtggtggtggtggtggt 679  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3051 gccggaacctgaaggtcagagtggtggtggtggtggtggtggtggtggtggtggtggt 3110  
QY 680 ggaagagacaggtagctcagcttctgagaggttataatgatacggatgattggtatca 739  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3111 ggaagagacacacagctcagcttctgagaggtcctatgacccacagagggatggtca 3170  
QY 740 tgggtgagatgagaatgacatcagagcttgaatggtggtggtggtggtggtggtggtggt 799  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3171 ggtgtgagcagagatattagacacataaagtattcctacggtgaatcatttgggtg 3230  
QY 800 tggtagtcagaagcgtgtttgttgcggagccacacatgaatgaataaagaattgagac 859  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3231 tggtagtcagaagcgtgtgttgcggagccacacatgaatgaataaagaattgagac 3290  
QY 860 gagatgatgtagctgataaagatgagagagcagcagaaggaagcaaatgcatagatt 919  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 3291 gtgaagaatgtaacatgataatgagagagcgtcgaaggaagcaatgcatagatt 3350  
QY 920 ttatcatgagatttctcataaataattatcaatggtatgaggggaaagagactcaatgta 979  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3351 ttatcatgaaactgcctcatalaatttgcacacctgttgcgagagagagggcccaagtgtga 3410  
QY 980 gtgagagcagagaagaagaggttgcgaattgtcgtggtggtggtggtggtggtggtggtggt 1039  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3411 gtggtgagcagagaagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3470  
QY 1040 tgattttagtaggctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1099  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 3471 tccgtcgtgatgagcgacgctcagccttgcagacagaaagcgagcagtggttcaggtcg 3530  
QY 1100 cacttgagaagcgagcaaaaggttcgactacatcgltgtaacaccgactttcacta 1159  
Db 3531 cctctgataagccagaaaagtcgcagcacacattgtgtaagctcactcttcttcacag 3590  
QY 1160 ttcgaaagtgcacattgtgactgcacctaaagatgaaatgctgcgtgcggaagagcgac 1219  
Db 3591 ttcgtaatgctgaactgcgtcgtgttccgataagtgaaatcattgtgcggaagaaac 3650  
QY 1220 atgctgaactaaatgycaaaacgagctcatalatattcactctgtgatacagagatata 1279  
Db 3651 atgataaactcctgaagaagagacatttacttcaactctgacaaatgcagacagcgag 3710  
QY 1280 aaaaagctgataacagatgagatcactatctcta-----ctga 1322  
Db 3711 gaattgaattgataagaaaatgcagctgataatccaaaagtgaattgtacgttcgtg 3770  
QY 1323 aagaagaaccactcactcctcgtcactctgtgaagagcatca----- 1366  
Db 3771 aaatgtcttccaatgtatcaagaatccagctctaataagaaaaagatcaactcgtagaagt 3830  
QY 1367 -----agtcagacttcattgcacaagctgcgaaatcccaatctaaagagatagtc 1420  
Db 3831 tccgtgatacacaagcccaagacagaagcttagtaccaaaagagctcgtgatgaagta 3890  
QY 1421 tccctgaagtctcctcatalaaattttaagttaaacaagcctgaatgccttcttggtg 1480  
Db 3891 taccctcaggttcccttcttgagagattatgaagctaaatttaactgtaatgyccttatttcg 3950  
QY 1481 tctctggagcatttgccttctgttcttaaatggaactgttcacagactttccacatcct 1540  
Db 3951 tctgtgtgtatatttgtgcacatataaatgtaggagcctgcacaacagacttgcataatat 4010  
QY 1541 tgcgaaaataatcaaccatglttgcgaaataatgataaaaaccaataaag---catgagt 1597  
Db 4011 ttccaagaattatagagggttttaccagaatttgatgatcctgtaaaaacagacagata 4070  
QY 1598 cagaaattatccatgatattccgtcatcttgggtgtatttggcttgcagtatttca 1657  
Db 4071 gtaactgttcttccactatgttctcagcccttgcgaattatcttcttattacatttcc 4130  
QY 1658 tgcaggaattatttcaacagacagaggaatttaacgatgagatlaagacacttgcg 1717  
Db 4131 ttcagagtttccactatgtgcagaaagctgtagagatcctccacaaagcgctccgataatag 4190  
QY 1718 ccttcaaaagccatgltatatactcagatattgctggttgcagtgaagaaagaaacagacag 1777  
Db 4191 tttctccgatacctgctccagacagatgltgagtgtgttgcagtccctaaaaacacacacgtg 4250  
QY 1778 gaggccttgacaacaatatattagccatagataatgcacaaattcaaggaacacaggttcca 1837  
Db 4251 gagcattgactccagcgctgcgccaatgactgtctcaagttlaaggggctataggttcca 4310  
QY 1838 ggaattgcgttcaacacacaaaatgcacccaacatgcagacttcgattatcattcttca 1897  
Db 4311 ggctgtgcttaattaccccgaaatataatgcacaaatcttgcgacaggaataatatactcttca 4370  
QY 1898 tataatgagtggagatgacatctccgtatctcgtagtatgcctcagtaactgcgcgtgacag 1957  
Db 4371 tctatgtgtgcgaactacaactgltactctttagcaattgtaaccacatctgcataagcag 4430  
QY 1958 gaattgataaacccgacgaatgcagactggttgcacaagaataagcaagaaacttaagc 2017  
Db 4431 gagctgtgtaaatgaaaatgltgtcgtcgcgaacagcacttgaagaataagaacactgaag 4490  
QY 2018 atgcttgagaagatagcaacttgaagcttgcgagataatactacatcagtgltcataa 2077  
Db 4491 gtgcgtggaagatcgtcactgaagcaataagaaactccgaacgcgtgttcttcttgacc 4550  
QY 2078 gggaaaaagccttcgagcaaatgtatgaaagatgcttcaagactcaacaagaataact 2137  
Db 4551 aggaagaagaagtgtgcacaatgtatgctcagagtttgcaggtacacataagaactctt 4610  
QY 2138 cgaagaagcacagatatttgaagcgtgtatgcatcagccaagccttataatattgt 2197  
Db 4611 tgaagaagacacacacttcttgaaatcaatttccctaccacagcaatgataatttct 4670  
QY 2198 cctatgcacagaggttctcgaatttgcagcctatttaattcaagcttgcagcaatacc 2257  
Db 4671 cctatgctgagtgttccgtgttgcagcctactctgtgtgacaataactatgactgttgc 4730  
QY 2258 aggcagtgtcatagtttcttactgcgaattgcataatgagctatgcctatgcgaagacgc 2317  
Db 4731 aggatgtcctgtatgatttcttcaagcgtgtcttgcgtgcagtcgcgtgcgagtcga 4790  
QY 2318 tegtgttgcctctgataatttccaagccaatccggggcgtgcagctgttgccttgc 2377  
Db 4791 gttcatlgtcctctgactatgcacaaataatcagagccacatcaatgata 4850  
QY 2378 tgaagaagaaccaataatagacagccgactgaagaaggaaggaagccagacagatgtg 2437  
Db 4851 tgaaaaaaaccttctgattgcagctacagccagcgaaagccttaatgcgaaacatttgc 4910  
QY 2438 aaggaatttagagtttcgagaagctccttcttcaatccatgctgcgccagatgtttca 2497  
Db 4911 aaggaatgttcaacattgtgtgaagttgtatccaactatccaccgcagccagatcccg 4970  
QY 2498 tctctcgtgtctatccctcagatgtgagcggaagaagagatgattgttgcggagca 2557  
Db 4971 tgccttaggaactcgaagcctgcgaggtgaagaaggccagagcgtgcgttgcgtgcagca 5030  
QY 2558 gcgcgtctggaagaaagcacttctcgttcaactcttcgacagactttagaaccccgctgaag 2617  
Db 5031 gtgcgtctggaagaaagcaacagtggttccagctccttcggaagcgttctacgaccttgcg 5090  
QY 2618 gacaagtgcctgttgatgtgtgagtcgaaaagaatgaaatgaaatgaaatgaaatgaaatg 2677  
Db 5091 ggaagaatgcgtcctgttgatgcgaaagaataaagagactgaaatgttccagtcctccgagc 5150  
QY 2678 aaatgaacatcgttccctcaagagcctgtgtcttcaacttgcagacttgcggaagaaatc 2737  
Db 5151 acctggacatcgtgtcccgagagccacatccctgttcttgcagcagatctgcggaagaaatc 5210  
QY 2738 cctatgtgacaacacagcgtgtgtgcattgaatagatgaaatgaagaagccgcaaatgcag 2797  
Db 5211 cctatggaagacaacacagcgtgtgttccaggaagaagatcgtgaaggccagcaagagag 5270  
QY 2798 caaatatcattcttattatgaaggtccctgcgagaatacaacacacaaagtgtgactga 2857  
Db 5271 ccacatacagcttccatcagctcagctcactgcttaataatagcactaaagttagagaca 5330  
QY 2858 aaggaacacagcttcttgcgagcgaagaacaaagactagttatgcaagggcctctcc 2917  
Db 5331 aaggaatccagctcctgtgtgcgaagaacaacagcattgcatalagctcgtgcctctgtta 5390  
QY 2918 aaaaaccacaaaatttattgttgcagtgagccacttccagccctcagataatgacagctgaga 2977  
Db 5391 gacagctcatalatttgcgttgcgttgcagtgagccagccagctcgtgataacgaaagagaa 5450  
QY 2978 aggtgttcaatgctcccttgcataaagccagcgaggaagagacatgctcagtgactc 3037  
Db 5451 aggtgttccaaagaagccctgcgaacaaagccaggaagggccgacactgcattgtatgtctc 5510  
QY 3038 acaggtcctctgcagatccgaagcagatgttgcagtgatgtgttcgcaaatggaagataa 3097  
Db 5511 accgctgttccacacatccaaatgcagacttaataatgagtggttccgaaatggaagatca 5570  
QY 3098 aggaacaaggaacttcaagagctcctgaaatctgagaacatataatttaagttagtga 3157  
Db 5571 aggaagatgtgcagcatcagcagctgtgcgcagagaagacatctattttcaatgtgta 5630  
QY 3158 atgcacagtcag 3169  
Db 5631 gttccagagctg 5642

RESULT	12
ID	AAZ49332 standard; cDNA; 3860 BP.
AC	AAZ49332;
DT	14-MAR-2000 (first entry)
DE	Human wild-type multidrug resistance-1 (MDR-1) cDNA.
KW	Multidrug resistance; MDR-1; P-glycoprotein; transmembrane efflux pump; haematopoietic stem cell; transduction; bone marrow transplantation; chemotherapy; radiation therapy; cancer; gene therapy; gene replacement; genetic defect; thalassemia; Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion; cytokine; wild-type; ds.
OS	Homo sapiens.
FT	Key
FT	CDS
FT	Location/Qualifiers 1..3843 /*tag= a /product= "Human wild-type MDR-1 protein" /replace (553..555, GGT) mutation /*tag= b /note= "cDNA sequence of G185V human mutant MDR-1 given in AAZ49333"
PN	WO9961589-A2.
PD	02-DEC-1999.
PF	27-MAY-1999; 99WO-US11825.
PR	28-MAY-1998; 98US-0086988.
PA	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL. Sorrentino B, Bunting K; WPI: 2000-072615/06. P-PSDB: AAY58186.
PT	Ex vivo expansion of hematopoietic stem cells transduced with a sequence encoding human multidrug resistance-1, used for bone marrow transplantation -
PS	Claim 10; Page 68-70; 113pp; English.
XX	This sequence represents cDNA encoding human wild-type multidrug resistance protein MDR-1. MDR-1 is a transmembrane efflux pump, responsible for the export of drugs from cells, particularly cancer cells. Wild-type MDR-1 shows increased resistance to etoposide and decreased resistance to vinca alkaloids compared with a mutant form (AAY58187) where the Gly at position 185 is replaced by Val. The invention relates to transducing haematopoietic stem cells with nucleic acid encoding an MDR protein and culturing the modified cells. The modified haematopoietic stem cells are useful in bone marrow transplantation (to reconstitute haematopoietic systems in patients who have undergone chemotherapy or radiation therapy) and in ex vivo gene therapy of genetic defects in cells derived from haematopoietic stem cells, e.g., thalassemia, Gaucher's disease, sickle cell anaemia or leukemia. The modifying cells can also be used to identify factors involved in regulating proliferation and differentiation in haematopoietic stem cells. Haematopoietic stem cells that express MDR-1 will be protected against chemotherapeutic agents, so can be engrafted while the patient is undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells provides sufficient cells to permit standard biochemical analysis. Overexpression of MDR-1 allows cytokine-driven expansion of haematopoietic stem cells by at least 10-fold compared with a maximum of 4-fold in known procedures.



DT	14-MAR-2000	(first entry)
XX	Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.	
DE		
XX	Multidrug resistance: MDR-1; P-glycoprotein;	
KW	transmembrane efflux pump; haematopoietic stem cell; transduction;	
KW	bone marrow transplantation; chemotherapy; radiation therapy; cancer;	
KW	gene therapy; gene replacement; genetic defect; thalassemia;	
KW	Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;	
KW	cytokine; mutant; ds.	
XX		
OS	Synthetic.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..3843
FT		/*tag= a
FT		/product= "Human G185V mutant MDR-1 protein"
PN	W09961589-A2.	
XX		
PD	02-DEC-1999.	
XX		
PF	27-MAY-1999;	99WO-US11825.
XX		
PR	28-MAY-1998;	98US-0086988.
XX		
PA	(SUJD-) ST JUDE CHILDREN'S RES HOSPITAL.	
PI	Sorrentino B, Bunting K;	
DR	WPI, 2000-072615/06.	
XX	P-PSDB: AAV58187.	
PT	Ex vivo expansion of hematopoietic stem cells transduced with a	
PT	sequence encoding human multidrug resistance-1, used for bone marrow	
PT	transplantation -	
XX		
PS	Example 1; Page 79-82; 113pp; English.	
XX		
CC	This sequence represents cDNA encoding human G185V mutant multidrug	
CC	resistance protein MDR-1, where the Gly residue at position 185	
CC	of the wild-type protein (AAV58186) is replaced by Val. MDR-1 is a	
CC	transmembrane efflux pump, responsible for the export of drugs from	
CC	cells, particularly cancer cells. The wild-type MDR-1 shows increased	
CC	resistance to etoposide and decreased resistance to vinca alkaloids	
CC	compared with the G185V mutant. The invention relates to transducing	
CC	haematopoietic stem cells with nucleic acid encoding an MDR protein	
CC	and culturing the modified cells. The modified haematopoietic stem	
CC	cells are useful in bone marrow transplantation (to reconstitute	
CC	haematopoietic systems in patients who have undergone chemotherapy or	
CC	radiation therapy) and in ex vivo gene therapy of genetic defects in	
CC	cells derived from haematopoietic stem cells, e.g., thalassemia,	
CC	Gaucher's disease, sickle cell anaemia or leukaemia. The modified	
CC	cells can also be used to identify factors involved in regulating	
CC	proliferation and differentiation in haematopoietic stem cells.	
CC	Haematopoietic stem cells that express MDR-1 will be protected against	
CC	chemotherapeutic agents, so can be engrafted while the patient is	
CC	undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells	
CC	provides sufficient cells to permit standard biochemical analysis.	
CC	Overexpression of MDR-1 allows cytokine-driven expansion of	
CC	haematopoietic stem cells by at least 10-fold compared with a maximum	
CC	of 4-fold in known procedures.	
XX		
XX		
SO	Sequence 3660 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;	
Query Match	35.3%; Score 1122.4; DB 21; Length 3860;	
Best Local Similarity	62.0%; Pred. No. 1.5e-293;	
Matches 1688; Conservative	0; Mismatches 1096; Indels 48; Gaps	
000	ttgttaaggtataacacgaatctcaagaatgcaagaatttggcataaaagactatag	259

Db	821	ttgnaaggtcaacaacaaaatttagaagaagcttaaaagaattgggtataagaagaagcttatca	880
Qy	260	cttcaaaaggtctctctgtgtgtcgtgtgtaactctttatataatgagaaactatgagactgtctt	319
Db	881	caagccaatttctataagtgctgtctgtcttccctgtcgtatctatgcaattctatgtccttcgctc	940
Qy	320	tttgtagtggaacctctctgtatcttcaatgagaaactggaataacatcggtgagctgtctt	379
Db	941	tcgtgataggacacacctgtctctctcagggga-----atatctatcttgacaaggtac	994
Qy	380	ttgtcgtttctctttagtgttaatccatagagaagtatgtgattggggagcagctccctcaact	439
Db	995	tcaactgtattcttctctgttatataatctggggcttctaagtggttggaagcgacatcccaagca	1054
Qy	440	ttgaaaccttcgcaataagccccggagagcgtccttccatattctccaggttaattgatataga	499
Db	1055	tttagaagcatttgcaaatatgcaagaagggagcgacttatgaaacttccaaagtaattgtatata	1114
Qy	500	aaccagatataagataacttttccacagctggaataaacctggaatccatagaaagaaactg	559
Db	1115	agccaagtatgtacagatattctgaagaagtgggcacaacaaacagatatatttaagggaatt	1174
Qy	560	tgtgaatttaaaaaagtcttcttctaataatccatccaaagcaatctcaaggtcttcaagaag	619
Db	1175	tgtgaattcaagaagtctcaacttcaattagtaaccctctccgaaaaagaattaaagctcttgaag	1234
Qy	620	gtctgaatctcaagaataaagtcctggagagaacagctcgccctgtgtcgtctcattatgagcagtg	679
Db	1235	gcttgaaacctgaaaggtgcagaagtggggcagaacggtggtcctgtgttggaacagctgctgtg	1294
Qy	680	ggaagagtagctgttagctccagctctcttcgaagggtatatagtatccgaatgaaatgagcttata	739
Db	1295	ggaagagtagcaacagctccagctcagctgtacagagctctatgataccccaagagggtgata	1354
Qy	740	tgtgtgtagtgagaatagacaatacagaagctttaaagtgtggtggcgatatatgagaaactatgag	799
Db	1355	gtgtgtgtagtgacaagagatatataagacaataaagttaaaggtttcttaagggaaatcatctgtgtg	1414
Qy	800	tgtgttagtgacaagagcgcgtttctgttcggggacccaatcgataacataatcagaattatgagac	859
Db	1415	tgtgtgtagtgacaagaaacctgtatgtgtttggccaacccaagctagctgtaaaaaccttcgtcatggcc	1474
Qy	860	gagatgtagtgactgtagtgaagagatgtggagagagcagaagggaaagcaaatgctgatatgact	919
Db	1475	gtgaaaaatgtcaccaatgtagagattgagaagaactgttcaaggaagaaacaaatgctcatgact	1534
Qy	920	ttatatatggaatttccctataataatttaatacatcttgtagggaaaaagaagctcaaatgga	979
Db	1535	ttatatatgaaaactctccctacataaattttgaacccctgtgtggagagagaggggcccagttgga	1594
Qy	980	gttgagagggcacaacaaacagagatctgcgaatttctgtgtccttaagtctcgaaaccccagaatttc	1039
Db	1595	gtgtgtgtgggcagagaagcagaagatctgcgcatttttgctgacagtgcctgtgttcgcaaccccagaatttc	1654
Qy	1040	tgaattttagatgaaggtactacgtctcgtcccttggaattccagaagaagcaagctgctttcaagctg	1099
Db	1655	tccctgtgatgtagagccacgtccacagcctcttggaacacagaagcagaacagatgtgttcaagtggtg	1174
Qy	1100	caacttgagaagaagcgagaacaaagtctggagcttcaaatctgttgtagacaacagcaattctactata	1159
Db	1715	cctcttgataaagtcacgaaaaaggtctggagcccaactgtgtatagatctcatgcttgttctacaag	1774
Qy	1160	ttcgaagatgtagaatattgttatctgaaccttaagaagtatggaatgtctgtgcggagaagaagcaac	1219
Db	1775	ttcgtgtaatgtagcagctcatcgtctgtgtttcgatgtagtgaggtcatctgttggaagaaggaatac	1834
Qy	1220	atgtctgaactaaatgtgcacaacagaggtcatataattatcaactctgtgagtcatcaagaatata	1279
Db	1835	atgtgtagaactcaatgaagaagaaagcaacttacttccaacactgtcacaatgacagagaacag	1894
Qy	1280	aaaaagctgtatgaacagatgtgagttcaatgacatatctta-----ctga	1322
Db	1895	gaaatgaaatttgaatttagaaaaatgcagctcatgtaatccaaaagtgaattgtatgtcctgtgtg	1954



[illegible]

OY	2378	ttgaaagaagaaacccaatatatgacaagccgcagatcagaagaagaaagaaacccaacacatcgtg	2437
Db	3035	ttgaaataaaaccccttctgtatgtaacagctacacagacgcggaagagcctaattgcccgaacacatctg	3094
OY	2438	aaggaaattttagagtttcctcgagaagctctctctctctacatccatgctccgaatgttttca	2497
Db	3055	aaggaaatgttcaacatttcttggtggaagtctgtattcacaattacccaccgcgaacccggaatccag	3154
OY	2498	tcctccgttggcttaccctcagatcttgagcagagaaagaaacagtaacatttgttggaggaca	2557
Db	3155	tgtctcagagagccttgagcctctgaggttgaaagaagggccagacgcttgctctgttggaggaca	3214
OY	2558	ggcgctcttgggaagaagacattcgtctcaactctctgaaagagatctatgaacccctgtcacg	2617
Db	3215	gttgcctcttgggaagaagacagctgtgtccacgctcctctggagcggttctacagccctctggag	3274
OY	2618	gacaagtgctcttcttgatctggtgtgagtcgaaagaatctgaattgaacgttggtccctccg	2677
Db	3275	ggaaagtgcttgcttgatctgagtcgaaagaataaagcgaattgaattgaatgtgtcttcgagcac	3354
OY	2678	aaatagacaatctgtcttcctcaagagccctgtgctctctcaacgtcgagcaatgtctgagaacatcg	2737
Db	3335	accgtggagatcgtgtctccacgagcccatcctgttcttgatcgtcagacatgtctgagaacatcg	3394
OY	2738	ccctatgtgtgacaacagacgctgtgtgtgcctatgatatgatacacaagaacgcgcaaatcag	2797
Db	3395	ccctatgtggagacaacagacggtgtgtgtcacaagaagaatctgttgaggcgcaagaaggg	3454
OY	2798	caaatatccattctcttatttgaaggtctccctctgagaaataacacacacaagtctgaccta	2857
Db	3455	ccaacatatacgtccttcatctgaagatcagctgtcttaataataatgacactaaagtaagagaca	3514
OY	2858	aaggagacacagctcttcctcgccgcgcagaaacaaagacatactatctcaagggtctctcc	2917
Db	3515	aaggaaatcagctctctctcggtgcgcagaaacaaacgcatctgcatagctctgtgcctctgta	3574
OY	2918	aaaaacccaataatttatgttctgatatgagcgcaatctcagccctcgataatgacagtga	2977
Db	3575	gacagccctcatatttcttcttcttgatgaagccacgcgtcactctgtgatacagaagtga	3634
OY	2978	agggtgttcacgacagccctctgataaagccagagccggaagagacatgagctagtggtcacct	3037
Db	3635	agggtgttcacgaagccctctgataaagccagaaagagccgcacctgacatgtgattgtctc	3694
OY	3038	acagcctctctgcaatctcagacaacgcagatttgaatgttctgtcacaatctggaagaagata	3097
Db	3695	accgcctgtccacatccatccagaaagcagacttaatgatgtgttccgaatgtgcagagca	3754
OY	3098	aggaaacaagaaactacacgaagctcctctgaaaatctcgaaacatataattttaagttagtga	3157
Db	3755	aggaaactgagacgcatcagcagctgtgtgcacagaaagatcatattttccaatgtgca	3814
OY	3158	atgcacagctcag	3169
Db	3815	gtgtccaggtcg	3826
RESULT_14			
ID	AAV6533	AAV6533 standard; cDNA; 4264 BP.	
AC	AAV6533;		
XX	08-JAN-1999	(first entry)	
DE	Mutated human P-glycoprotein gene designated mdr-delta-F335.		
XX			
KW	mdr1-delta-F335; cyclosporin-resistant deletion mutant; human;		
KM	P-glycoprotein; cyclosporin; drug efflux;		
XX	modulator-resistant multidrug resistance; cancer therapy; ss.		
OS	Homo sapiens.		
XX			







Db 959 ttgaaagtacacaaaaatttgaagaagcctaaagaattcgggataagaagaatatta 1018  
Qy 260 cttaaaagtctctctgtgtgtgtgtctcttattgaatggaaacctatgtctgt 319  
Db 1019 cagccaattatcttaagtgctgtcttcctcgtatctatgaatcatctatctcgtgct 1078  
Qy 320 ttgtgtatggaaacctctgtatctcttaatggagaacctggatatccatcggaactgttc 379  
Db 1079 tctgttatgggaacctctgtctcctcgaaggga-----atatctcttggacaagtac 1132  
Qy 380 ttgtcttctctttagtgatlaatccatagcagttatgtcatgtggagcagctccctac 439  
Db 1133 tcaactgtatcttctctgtatlaattggggcttttagtgttggacagcatcccaagca 1192  
Qy 440 ttgaaaccttgcgaatagccggagagctgccttcatattttccagttatttgaaga 499  
Db 1193 ttgaagcatttgcgaatgaagaagagcagcttaagaaaccttcaagaatattgaata 1252  
Qy 500 aaccagatatagaatacttccacagctggatataaaccttgaatccatagaagaacty 559  
Db 1253 agccaagatattgacagcattccgaagagtgggcacaaaccagataatatatgaaggaaatt 1312  
Qy 560 ttgaaattaaaaatgttcttcaatlaatccatcaagaaccatctatcaagattctgaag 619  
Db 1313 tggaaattcagaatgttcaacttcaagttacccatctcgaaagaagttaagattctgaag 1372  
Qy 620 gttctgaactcctgaattagctctgtgagagacagctgccttggctgtctcaatgtcagtg 679  
Db 1373 gctctgaaccttgaaagtgagagtgggcgacggtggccctgggttggaaacaaatgtcgtg 1432  
Qy 680 ggaagaagtagcgtatgcagctctcgagaggtatcatgcatcggatgatgtcttatca 739  
Db 1433 ggaagaagacaacaatcgaactcagctgagagaggtctatgaaccacagaggagtgatca 1492  
Qy 740 tggtgagatgagaatgaatcagacttcaaatgtgcggcatatccagaaacctatgtgag 799  
Db 1493 gtgttgatggagaagatatltgagaccataaagtlaaggtttctaagggaataatcatgtg 1552  
Qy 800 tggttatgacaagaagccttctgttcggagccacatcaagtaaatccaaatcatgtgac 859  
Db 1553 tggtgagtagcagaacccgtgaattgtttgcacacagatagctgtgaataacatctgtatggcc 1612  
Qy 860 gagaatgaatgtactgtatgaagaatgtgagagagcagcagaagggaagcaaatgtgata 919  
Db 1613 gtgaataatgtccacatgtgagatltggagaagctgtlcaaggaaagccaaltgctatgct 1672  
Qy 920 ttatactgtgagttccctaataaatttaatacatgtgttaggggaaaaagagctcaaatga 979  
Db 1673 ttatactgaactcgtcatataaatttgaacacctgtgttgagagagaggggcccagttga 1732  
Qy 980 gtggagggcagaaaacagagagctcgcaattgtcgtgaccttagttgaaccccccaagattc 1039  
Db 1733 gtgggtggcgagaagaagagatgcgcacatgtgcagctggcttgcgaaccccccaagattc 1792  
Qy 1040 tgaattttagatgaagctcgtctgcgcctgtgattcgaagaagcagatcagctgttcaagcty 1099  
Db 1793 tccgtctgtgattgagggccacgtcagccttggacacagaaaagcgaagcagtggttcaggtg 1852  
Qy 1100 caatgtgagaagcgagcaaaagctcggaactcaaatcgtgtgttagcacacagactttcta 1159  
Db 1853 cctctgtataaggccagaaaagctcgacacacatgtgtgtagcatcgtttgtctacag 1912  
Qy 1160 ttcgaagatgcagatttgaattgtgacccctaaagatgtgaatgtgtgcggagaagaagcac 1219  
Db 1913 ttctgtatgtcgtgacgtcgtgttccgaatgtgagatcatltgtggaagaaatcaatc 1972  
Qy 1220 atgcgaactaattggcaaaaaggtctcatattcatcacttggatgtgcacagagtatca 1279  
Db 1973 atgatgaactccttgaaaagaaagaaacttacttcaaaacttgcacatgtcgacagcag 2032  
Qy 1280 aaaaagtgtatgaacagatgtgagtcgaatgtacatctc-----ctga 1322

Db 2033 gaatgaagttgaattagaataatgcagctgtatgaatcccaaaagtgaatltgactgtg 2092  
Qy 1323 aagaagaagcaactcactctctcgtactctgtgaagaagctca----- 1366  
Db 2093 aaatgtctcaaatgatccaagatccagttcaataaagaataaagaataactcgttagagtg 2152  
Qy 1367 -----agtcagacttcatltgacaaggttgagaatcccaaccaatctaaagataagtc 1420  
Db 2153 tccgtgtatccaagcccaagacagaagaagcttgaatccaagaagagctcgtgatgaagta 2212  
Qy 1421 ttcttgaagctctctatataaattttaaagttaaagaacccctgaagagccttttgg 1480  
Db 2213 taacctccagttcccttggagagatlaagctaaatlttaactgaatgtgaccttatttg 2272  
Qy 1481 ttctgggacatttggctctgttccaaatggaaactgttcatccagatlttccatcatct 1540  
Db 2273 ttgtgtgtgtatttctgtccattataaattgagagcctgtgcaaccgcatttgcataat 2332  
Qy 1541 ttgcaaaaattataacacatgttggaaatlaatgaataaacacatlaag--gatagt 1597  
Db 2333 ttccaagaattatagggtttttaacaagaatttgaatccctgaacaacaacagcagaata 2392  
Qy 1598 cagaattatccatgaatctgtcatcttgggtgttatttgccttgcgtatttca 1657  
Db 2393 gtaacttcttcaactatgttctagcccttggaaattatlttcttattatcatlttcc 2452  
Qy 1658 ttgcaagattattttacggcagagaggaatttgaagatgaatgaagacacttg 1717  
Db 2453 ttcaaggtttcaacttggcaaaagcttggagabcttccaacagcgctccgatacag 2512  
Qy 1718 ccttcaaaagcattatatacagatatgtcgtgttgaatgaagaaggaacagcag 1777  
Db 2513 ttcttcgactcagtgctcagacagagtgagttgtgttgaatccctaaacacacacty 2572  
Qy 1778 gaggtcttcaacaataatlagccatagatatagcaacaatccaaaggaacaggttca 1837  
Db 2573 gagcatltgactcagcagctcggcacaatgtcgtccaaagttlaaagggtctatagttcca 2632  
Qy 1838 ggaatggcgtcttaacacaaaatgtcaactaaatgagacttcaagttatcatcttcc 1897  
Db 2633 ggtctgcgttaattatcccaagaatatagaatcctgggaacgaatataatcttcc 2692  
Qy 1898 tatatgtatggagatgaatcactcgtgattctgagatgtgtccagttacgttcggtgacag 1957  
Db 2693 tctatgtttggaactaaacatcagttactccttagcattgttccctcatatagag 2752  
Qy 1958 gaattattgaaacccgacgaatagactgtgattggcaaaaagatatagaagaacttaagc 2017  
Db 2753 gaattgttgaataatgtgtctgtgacaagcagctgaagaatataagaagaactagag 2812  
Qy 2018 atgcctggaagaatgaactgtgaagcttggagaatatacgttactatgttcaataca 2077  
Db 2813 gtgcctgggaagctcgtcactgaagcaatagaaacttcgaaacggttcttcttgactc 2872  
Qy 2078 gggaaaaagccttcgagcaaatgtatgaagagatgcttcaagctcaacaganaatcct 2137  
Db 2873 agagagcagaagttgacaatatatgtcagagatttgaagttaccatacagaanaactct 2932  
Qy 2138 cgaagaagaacagaattatttgaagcgttattatgacttcaagcatttatatatatttg 2197  
Db 2933 tgaagaaagcacacacttcttgaattatcatcttccatcccaaggaatltgattatct 2992  
Qy 2198 cctatgcagcagaggttctgaatttggagcctattatccaagcttgaagaaatgacccag 2257  
Db 2993 cctatgcgtgagatgttccggttggagcctacttgggtggcaataaactcatgaacttg 3052  
Qy 2258 aggcagtttcatagttttacttgaatgtcatatggaatgtgacatcggaaaacgc 2317  
Db 3053 aggatgttctgttagtatatttcagctgtgtcttgggtgcagtgccgttgggagca 3112  
Qy 2318 tctgttctgctcctgaatatccaaagcgaatcgggggctgtgcgactctgttgcctgt 2377  
Db 3113 gtctatttgcctcgtactatgccaagccaataatcagcagccacatcatcatgtatca 3172



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:46:40 ; Search time 14771.4 Seconds  
(Without alignments)  
3548.174 Million cell updates/sec

Title: US-09-873-409-12  
Perfect score: 3177  
Sequence: 1 atgcacatcattgacag.....atgcacagtcagtcagtcga 3177

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.om:\*  
20: em.or:\*  
21: em.ov:\*  
22: em.pat:\*  
23: em.ph:\*  
24: em.pl:\*  
25: em.ro:\*  
26: em.sts:\*  
27: em.sy:\*  
28: em.un:\*  
29: em.vi:\*  
30: em.htgo\_hum:\*  
31: em.htgo\_inv:\*  
32: em.htgo\_rod:\*  
33: em.htg\_hum:\*  
34: em.htg\_inv:\*  
35: em.htg\_rod:\*  
36: em.htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1130	35.6	4279	6 AX105057	AX105057 Sequence
2	1130	35.6	4279	6 AX105078	AX105078 Sequence
3	1130	35.6	4279	6 AX105080	AX105080 Sequence
4	1130	35.6	4279	6 AX105082	AX105082 Sequence
5	1126.8	35.5	4045	12 AF269224	AF269224 Synthetic
6	1126.8	35.5	4317	4 AF045016	AF045016 Canis fam
7	1126.8	35.5	4317	6 AX105059	AX105059 Sequence
8	1124	35.4	4646	6 HUMMDR1	149610 Sequence 2
9	1124	35.4	4669	9 HUMMDR1	M14758 Homo sapien
10	1124	35.4	4669	9 HUMMDR1	108557 Sequence 3
11	1124	35.4	6505	6 AR028671	AR028671 Sequence
12	1122.4	35.3	4264	6 AR028672	AR028672 Sequence
13	1122.4	35.3	4264	6 AR051647	AR051647 Sequence
14	1122.4	35.3	4264	6 AR051650	AR051650 Sequence
15	1122.4	35.3	4378	6 E02326	E02326 Multidrug r
16	1122.4	35.3	8630	6 AX012321	AX012321 Sequence
17	1120.8	35.3	8630	6 AX012320	AX012320 Sequence
18	1119.2	35.2	4669	6 AR055785	AR055785 Sequence
19	1114.4	35.1	4669	6 AR091275	AR091275 Sequence
20	1112.8	35.0	4186	6 AX108654	AX108654 Sequence
21	1112.8	35.0	4195	6 AX108656	AX108656 Sequence
22	1110.6	35.0	4018	5 GGA9799	AX009799 Gallus ga
23	1108	34.9	3988	6 AX024454	AX024454 Sequence
24	1108	34.9	4192	9 AF016535	AF016535 Homo sapi
25	1104.8	34.8	4296	10 CR0PGP1	M60040 C.grius p
26	1103.2	34.7	3987	10 CR0PGP1185	M59253 Chinese ham
27	1103.2	34.7	4304	10 CR0PGP1165	AX024455 Sequence
28	1097	34.5	3924	6 AX024455	M23234 Human membr
29	1097	34.5	3924	9 HUMMDR3	M60041 C.grius p
30	1095	34.5	4280	10 CR0PGP11	U78609 Ovis aries
31	1074.4	33.8	3858	4 OA078609	J03398 Mouse mdr g
32	1055.4	33.0	4084	10 MUSMDRA	AF257746 Rattus no
33	1047.2	32.9	4927	10 AF286167	AF286167 Rattus no
34	1044.6	32.9	4323	10 AF286167	M14757 Mouse mult
35	1043.4	32.8	4298	10 MUSMDR	AR123273 Sequence
36	1040.6	32.8	4233	6 AR123273	M60042 C.grius p
37	1039.4	32.7	3905	10 CR0PGP11	M24417 Mouse phosph
38	1028	32.4	3682	10 MUSPGLY1A	M33581 Mouse p-gly
39	1028	32.4	4924	10 MUSMDRIA	M30697 Mouse mult
40	1026.4	32.3	4356	10 MUSMDRAA	U17608 Xenopus lae
41	1024.2	32.2	4390	5 XU117608	AB029153 Felis cat
42	1023.4	32.2	3489	4 AB029153	M81835 Rat mdr mRN
43	1013.4	31.9	4254	10 RATMDRM	L15079 Rat p-glyco
44	1007.2	31.7	3912	10 RATPGLYCO	I33621 Sequence 1
45	856	26.9	2726	6 I33621	

## ALIGNMENTS

RESULT 1  
AX105057  
LOCUS AX105057  
DEFINITION Sequence 1 from Patent WO0123540.  
ACCESSION AX105057  
VERSION AX105057.1 GI:13921209  
KEYWORDS  
SOURCE  
ORGANISM  
dog.  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 4279)  
Stocker,P.J., Steinel-Crespi,D.F., Crespi,C.L., Relf,T.C. and  
Patten,C.J.  
TITLE  
P-glycoproteins and uses thereof  
JOURNAL  
Patent: WO 0123540-A 1 05-Apr-2001;  
FEATURES  
LOCATION/Qualifiers  
1..4279  
/organism="Canis familiaris"

```

CDS
/db_xref="taxon:9615"
17..3862
/locus="unnamed protein product"
/codon_start=1
/db_xref="GI:13921210"
/protein_id="CAC37718.1"

Query Match      35.6%  Score 1130: DB 6: Length 4279;
Best Local Similarity 62.4%: Pred. No. 2.7e-222;
Matches 1879; Conservative 0; Mismatches 1085; Indels 48; Gaps 5;

BASE COUNT      1294 a      834 c      1008 g      1143 t
ORIGIN
200 ttgtaaggtatcacagaatcacaagaatgacaagaatttgacataaaaggagactaag 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
843 TTGAAGAGTACAAACAAATTTAGACAGCTAAAGAAATGGAGTAAAGAAAGCTATCA 902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 cttaaaagtgcctcctggtgctgctgactcttatagaatgagaactagactgctt 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
903 CGGCAACATTTCTATTGTCGCCCTTCTTATGATCATGACATCATGCTGCTGCTT 962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
320 ttggtatgagaacctctgacttcttaatgagaacctgagatcacactgagactgct 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
963 TCTGTATGAGGACCTCTGCTGCTCTCCAG-----TGAATATTTCTATTGACAGATAC 1016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 ttgctgttcttctttagtgaatcatcatgacttcatgacagagcagctccactac 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1017 TCACTGCTCTCTTTTCGTATTAATTTAGGCGCTTTAGATTTGACAGACACCCCAAGCA 1076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 ttgaaccttcgcaatgcccagagagctgctcttcatatttccagagttatgataa 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1077 TTGAAGCATTTGCAAGCAAGAGACAGACAGCTTATGAATCTTCAAGATTAATTCATAATA 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
500 aaccagtatagataactttccacagctgatatgaacctgaatccatcatgaagaactg 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1137 AACCAACCATTTGACAGCTATTGCAAGAGTGACATTAACCCAGATTAATTTAAGGAAATT 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
560 ttgaattaaagaattcttcttcaatcatcatcaagaacctcatcaagaatttgaag 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1197 TGGATTTCAAAAATGTTCTTCACTTACCTTCTCGAAAAAGAGTTAAAGTCTTTAAAG 1256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
620 gtctgaatctcagaatlaagctcgtgagagacagctgctgctgctcgtctcaatg 679
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1257 GTCTCAACCTGAAGCTTACAGAGTGACAGACAGTGGCGCTGTTGGGAACAGTGGCTGCG 1316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
680 ggaagagtagcgtatgcagcttctgcaagaggtatatgacggagatgagcttata 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1317 GGAAGAGCAGCAGCCGCTGATGACAGAGGCTCTATGACCCCAACAGATGGCATGGCT 1376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

740 tggatgagtgaaatgacatcagagctttaaagtgcgagctatcagagaccatattgag 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1377 GTATGTGACAGACAGCATATTGACCATTAATGTAAGGCTCTTCTGGGAATTAATGCTG 1436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
800 tggattgtaagagccctgtttgttcggggccacccatcagtaacatataagtaggac 859
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1437 TGGAGATCGAGGACCGCTGTTGTTGCCACACAGATAGCTGAACATTCGCTATGGCC 1496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
860 gagaatgtagcagatgagaagatgagagagcaagcaagaagagcaaatgctatgagt 919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1497 GCGAAAATGTCACCATGAGATGAGATTTAGAAAGCTGTTAAGAAAGCAAGCTATGATT 1556
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
920 ttatcatgagagttctcctaataatlttaacatctgtaggggaaaaagagctcaatga 979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1557 TTATCATGAAGAACTCAATAAATTTGACACTGTTGGAGAGAGAGGGCCACAGCTGA 1616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
980 gtggaagagcagaagaagagagatcgcaattgctgctgcttaagtcgaaccccaagatc 1639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1617 GTGTGAGACAGAAACAGAGATGCGCATTTGCTGGGCGCTGTTGCGAACCCAGATTTC 1676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1040 tgaattagaatgagagctgctgctgctgattcagaagaagcaagtcagctgccaagctg 1099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1677 TTCTGCTGATGAGAGCAAGCTCAAGCTTGACACTGAAAGTGAAGCAGTGGTTCAGGTG 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1100 cactggaagagcgagcaaaagctgagactaacaactcgtggtagaacacagacttacta 1159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1737 CCTGAGTAAGGCGCAGAAAGCGCGGACTACCATTTGATGATGATCATCGTTTCTACAG 1796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1160 ttggaatgcaagattgattgtgtagacctaaagatggaagtcgagcgaggaagaagcag 1219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1797 TTGCTAATGCGGATGATGATGCTGCTGCTGTTGATGATGATGATGATGATGATGATG 1856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1220 atgctgaactaatgcaaaaacgagctatataatlttcaactgtgagtgcacagataat 1279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1857 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1916
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1280 aaaaagcttagta-acagatgagatcaatgacatattctctgaagaagaagacactca 1338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1917 GAAATGAAATTTGAGTGAAGAAATCCACTGCTGATGATGATGATGATGATGATGATG 1976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1339 ctctctctgactctgtgaagagc-----atcaagtcagacttcaatgcaagagctg 1390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1977 AAATGCTCCAAAGATTTGAGGCTGACGTTTAAATTAAGATCAATCTGCCAGAGTA 2036
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2037 TACATGACACCAAGGCGCAGACAGAGCTTGCTACAAAAGAGACTTGATGATGATG 2096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1423 --cctgaagctctctatlaaaaaatttaaaaglttaaaagcctgaatgagcttctg 1480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2097 TACCTCCAGTTTCTCTGAGGATTTCTGAAGCTGAACTCAAGATGAGTGGCTTATTTTG 2156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1481 ttctgggagacattgagctctgcttcttaaatgaaactgctcatcagatatttccatc 1540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2157 TGGTTGATATTTTGTGCTATTAATTAAGCGAGGCTGCAACGCACTTTCAATTAATAT 2216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1541 ttgcaaaatataaacattgctt---gaaataatgataaaacacataaagcatgag 1597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2217 TTTCAAGATTAATAGGATCTTTACCCGAGATGAGATCTCTGAAACGAAACGACGATA 2276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1598 cagaattatcatcatgatatctgcatatttggtgttatttgccttgcgaattattca 1657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2277 GTACATGTTTCTGTAATTTGTTCTAGTCTGGAATTAATTTCTTATTAATTTTCC 2316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1658 tgcagagattatttcaagcgagagcaggggaaattttaaagatgagattgaagactg 1717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2337 TCCAGGCTTCCACTTTGCGCAAAAGCTGGGAGATCCCTCACTAACCGCTTGCATACATG 2396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1718 cctcaagacatgttatcatagatatctgctggttgcattgataaagaagaagaagcag 1777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2397 TTTTCAGATTCATGCTGAGACAGATGCTGCTGTTTGAATGACCTTAAACACCACTG 2456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1778 gagcttgacacaatatatgacatatagacacaattcaagagcaagcaggttcca 1837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db	2457	GAGCATTTGACACACGAGCTTGCCAAATGATGGCGCTCAAGTTAAAGGGCTATAGGTTCA	1897
Oy	1838	ggaattgcgcctttaacccaataatgcaactaacaatggaacttccagttatcaatttcctta	1897
Db	2517	GGCTTGCTGTCATATTACCCACAATATATAGCAAACTTGGGACAGGACTTATTTATTCCTTAA	2576
Oy	1898	tataatgagtgaggagatgaacatccctgcatccgattatggtctccagtlactgycgtgacag	1957
Db	2577	TCTATGTTGGCATTTAACTTTACTCTGTAGCAATTTATACCCATCACTTTGCATATAGACG	2636
Oy	1958	gaatgatgtgaaccgcgcagcaatgacatgcatgatttgcacaagaatagataagaacttaagc	2017
Db	2637	GAGTGTGTTGAAATGAAAATGTTGTCTGGACAAAGCAGTGAATAATTAAGAAAGCTAGAG	2696
Oy	2018	atgcttgaaagaatagcaactgaaacttggagaataatagtctaatagttcatatgaacaa	2077
Db	2697	GAGCTGGGAAGATTGCTACAGAAAGCCATGAAAACCTCGCAACTGTGTTCTTTGACTC	2756
Oy	2078	gggaaaaagccttcogagcaaatglatgaagagatgcttcagatctcaacacagaataact	2137
Db	2757	GGGAGCGAAGATTGTAATACATGTATAGCACAAGTTTGCAATACCATTAAGAAACTCTT	2816
Oy	2138	cgaagaagaacagatattatggaagctgttatgcatctgaagccatgctttatatattt	2197
Db	2817	TGAGAGAAAGACACACATCTTCGGGGCTCATTTTCTATPACCCAGGCAATGATGATTTT	2876
Oy	2198	ccctatcagaagaagtctcgaatttggagccatttaatactcaagcttgaagaaatgaccag	2257
Db	2877	CCATATGCTGGCGCTGTTTCCGGTTTGGTGCTACTTGGTGGCAATATAGTTCAATGAACTTTC	2936
Oy	2258	agggcatgttcatatagtttcttactgtcaattgcatatagagctatgagccatcgaaaaaacg	2317
Db	2937	AGGATGTTCTTTTGGTATPCTCAGCTATTGTCTTTTGGTCCATGGCAGAGGGGCGACGTCA	2996
Oy	2318	tcgatttggtctccgaaatctccaagaacccaatcggyggctgcgcatctgtttgcctgt	2377
Db	2997	GTTTCATTTGGCTCCCTGACTATAGCCAAAGCCAAAGTATCACACACCACGACGTCAATGATCA	3056
Oy	2378	tggaaaaagaacccaatatagcagccgcgaatcaagaaggaaggaagacagacacatgt	2437
Db	3057	TTGAAAAAAGCCCTCTGATTTGACAGCTACAGCCCTCACGGCTCTCAAGCCCAATATACGTTGG	3116
Oy	2438	aaggaaatttagagtcttcgagaagctctctctctatcatcagtcgcccagaatgtttca	2497
Db	3117	AAGAAATGTGACATTTAATAGAGTGCTGTCTACATGCAATGCCACTGACACAGACATCCCG	3176
Oy	2498	tcctccgtgacttatccctcagtatctgagcgaagaaacagtagaatgtgvgggagca	2557
Db	3177	TGCTACAGGGGCTGAGGCTCTGAGGTGAAGAAAGGGCCAGACGCTGCGCCCTGTAGGTAGCA	3236
Oy	2558	ggcgctctgggaaaaagaactctgtctcaactctcgagagacttataaacccgylgcaag	2617
Db	3237	GTGGCTGTGGGAAGACACAGTTGTTTCAGTCTCTTAGAGGCTTCTATAGACCCCTTGGCGTG	3296
Oy	2618	gacaagtgcgttltgatggtgtgtagatgcaaaaagaatttgatgtacagctgagctcggtlccc	2677
Db	3297	GTTCAGTGTCAATTGATGAGGCAAGAGATTAAGCACTGATGTCCAGTGTCCGAGACAC	3356
Oy	2678	aaataagcaatcgttctcctcaagagcctgtgctctcaacttgcaagctgtctgagaacatcg	2737
Db	3357	ACCTGGGCAATCGTGTCTCAGAGACCCCATCTGTTTGTAGCTGCGAGCAATTCGAGAACTTG	3416
Oy	2738	ccctatgtgacacaagcccggtgtgtgtgcctatagatgagatcaagaagccgcaaatgcag	2797
Db	3417	CTTATGGAGACAAACGCCGGGTGTATTCACATGAAGAGATTTCGAGGCAACCCAAAGAGG	3476
Oy	2798	caaatatcccatctcttatgtgaagctctcccgtagagaaatatacaacacagaatltgactga	2857
Db	3477	CCAACTATACACCACTTCATGAGACACTCCCTGTGAGAAATACAAACCAAGATAGAGACGA	3536
Oy	2858	aaggagcacagcttctgctgcgcagaaaaagaactagctatgtgcaaggctctctcc	2917

BASE COUNT	1295 a	833 c	1008 g	1143 t
Db	3537	AAGGAACCCAGCTCTCTGTGTGGCCAGAAACAGCCATTGCCATAGCTGCCCTCTTGTTA		3536
QY	2918	aaaaaaccacaataattatctgtgtgaatgagagccacttcagccctcgataatgacagtga		2977
Db	3597	GACACGCTCATATTTTGGCTTTTGATGTAAGGTCACATGAGCTCTGGATACAGAAAGTGA		3656
QY	2978	agggtgttcacgaatgccttgcataaaccaggaacggaagacatgacctagtgtgctactc		3037
Db	3657	AGGTGTGTCACAGAAACCCCTGGACAAAGCCAGAGAGGCCGACCTGCTTTGTGATCGCCC		3716
QY	3038	acagcctctgtcgaatttgaagaacgaggtttgtagtgtgtctcgtgacaaatggaagataa		3097
Db	3717	ACCGCTTGTCACACATCCAGAAATGAGATTATATGTGTGTGTGATGAGTCAAGCA		3776
QY	3098	aggaaacaagaactatcaagaagcctccgagagaatcagacatatatttaagttagtga		3157
Db	3777	AGGACATGCGCAGCACATCATCAGACGCTGCTGGCCCAAAAGGATCATATTTTCCATGCTCA		3836
QY	3158	atgcacagtcag	3169	
Db	3837	GTGTCCAGGCTG	3848	
RESULT	2			
LOCUS	AX105078	4279 bp	DNA	PAT
DEFINITION	Sequence 22 from Patent WO0123540.			30-Apr-2001
ACCESSION	AX105078			
VERSION	AX105078.1	GI:13921228		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
SOURCE				
CDS				





Oy	2258	aaggcagatcctaatagtttllctacgcgaattgcatactgagagctatctgcatctcggaanaacgc	2317
Db	2937	AGGATGTTCTTTGGTATTTCTCAGCTATTGTCCTTGGTCCATGGCACTGGGGCAGGTCA	2996
Oy	2318	tcgattctggtccctcgaaatctccaaagccaaatcggggctctcgcatctglttgcctgtc	2377
Db	2997	GTTTCATTTCCTCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACTGTCATCATGATCA	3056
Oy	2378	tggaaagaaacaacaaatataagacagccgcagctcaagaagaaagcaagcaacatctg	2437
Db	3057	TTTGAAAAAACCCCTGATGTAGACAGCTACAGCCCTCACGGCCTCAAGCCAATAGTTGG	3116
Oy	2438	aagggaattagagtttctcgagaagctctcttctctatccatctgtcccaagatlttca	2497
Db	3117	AAGGAATATGTACATTATTAATGAGCTGTGTTCACCTATCCACTGCAACATATCCCG	3176
Oy	2498	tcctcccgctgagctlatccctccagatctgagcggagaagaacagtaagcttcttgaggaa	2557
Db	3177	TGCTCCAGGGGGCTGAGACCTTGAGCTATGAAAGGCCAGACGTGGCCCTGTAAGTACGA	3236
Oy	2558	gcggctgtgggaaagaactctgttcaactctcgagagacttataagcccgctgcga	2617
Db	3237	GTCGCTGTGGGAAGACACAGTTGTTCACTCCTGAGAGCGCTTATGACCCCTTGCGTG	3296
Oy	2618	gacaaagctcgttttagtctgtgtgatagcacaaagaattgaatgtatcaagtgtccttcc	2677
Db	3297	GTTTCAGTGCTAATTATGATGGCAAGACATTAACACCTGATATATCCAGTGCTCGAGAC	3356
Oy	2678	aaatagaatctgttctccaaagacccctgtccttcaactgcagagaaatctgtagaacatcg	2737
Db	3357	ACCTGGGCATCTGCTCTCAGAGACCCATCTGTTGACTGCGACATTTGCGAGAACATTTG	3416
Oy	2738	ccctatgtgacaaacagccgtgtgtgtgcattagatgagatcaaaagaacccgcaaatcg	2797
Db	3417	CCGTATGAGACAAACAGCCGGGCTGTTATCATAGTAAGAGATTATGCAAGCACCAGGAGG	3476
Oy	2798	caaatatcaattcttattatgtgaagctctccctcgagaataacacacaaagttagctga	2857
Db	3477	CCAATCATACACCACTTCATCGAACACTCTCCGAGAAATACACCAAGATAGAGACA	3536
Oy	2858	aagagacacagcttctctgagcgcacaaacaaagaactagctatctgcaagagctctctcc	2917
Db	3537	AAGGAACCAACGCTCTCTGGTGGCCAAACACGCAATTGCCATAGCTGGCGCTCTTGTTA	3596
Oy	2918	aaaaaaccaaaatttattgttggatgagccacttccagccctcgataatgacagtga	2977
Db	3597	GACAGCCTCATATTTTGGCTTTTGGATGAGACTACATCACCTCTGGATACAGAAAGTAA	3656
Oy	2978	agtgagtctgaagctgcctctgtataaagccagaagcgggaagacaaagctagtgtctac	3037
Db	3657	AGGTTGTCCAGAAAGCCCTGGACAAAGCCAGAGAGGCCGCGCACCTTGATGTGGCC	3716
Oy	3038	acaaagctctctgaattcagaacgcagattgatagtgtctgcacaatggaagataa	3097
Db	3717	ACCGCTTGTCCACCATCCAAATGCAAGATTATATAGTGTGTTTCAGAAATGGCAAAGTCA	3776
Oy	3098	aggaacaaggaactcatcaagagctctctgagaatcgaaatcgacatatalttaagttagtga	3157
Db	3777	AGGAGCATGGCACACATCAACACACTGCTGGCCAGAAAGCATCTATATTTTCCATGGTCA	3836
Oy	3158	atgcacagctcaag	3169
Db	3837	GTTGTCAGAGCTG	3848

	RESULT						
		3					
AXI05080							
LOCUS	AXI05080	4279 bp	DNA	PAT			30-APR-2001
DEFINITION	Sequence	24 From Patent WO0123540.					
ACCESSION	AXI05080						
VERSION	AXI05080.1	GI:13921230					
KEYWORDS	.						

Source	Organism	dog
REFERENCE	Canis familiaris	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.	
TITLE	1 (bases 1 to 4279)	
JOURNAL	Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Relf,T.C. and Patten,C.J.	
FEATURES	P-glycoproteins and 24 others thereof	
source	Patent: WO 0123540-A 24 05-APR-2001; GENTEST CORPORATION (US)	
CDS	location/Qualifiers	
	1..4279	
	/organism="Canis familiaris"	
	/db_xref="taxon:9615"	
	17..3862	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAC37722.1"	
	/db_xref="gi:13921231"	
	/translation="MDPEGGRGSAEKNFMKGRKKKKKKKKPKVSTFAMFYSN	
	WLDLVMYGMATMIIHGALPLMLVGNMNTSEFANAIGISRNKTPVITINESTNN	
	OHLFNHLEEMPTTAYYYSGAGVIAAYOVSEFNLGALRQFLTKRQEPHAIMQ	
	ETGFVDHNGELNTRITLDYSKINTEGIDKIGTFQFOSITPFIPTGIVFTGRKRL	
	VIIAISPVLGSAIMAKILSSFTDKELIYAKGAVAEEVLAIPIVIAFGQKEL	
	ERYKNMLEAGAGIGIKKAIYANISIGAFLIYASVAFYGTSLVSESYIGVU	
	TVFSEVLIGAFSIGQASPSIEAFANAGAAIEIKIIDNKSIDSKSGHPNIGK	
	NLEKRNHYSFSPSRKRYKILKGLNKKVSGQGVFAVNSNGKSTYQLORLDPPD	
	GMVQIDQDIFITINVRHLRETVGSQEPULFATTIENIRYGVENTMDEIKAAVE	
	ANATDITKLDPKFDITLGERGAOLSGQKORIAIARALVRNPKTILLDETSALDE	
	SEAVVAVALDEARKGRITIVIAHLSTVRNADVIAGFDGVYERKGNHDELMKGIY	
	FKILVMTQTRNEIILENATGESKESDLEMPKDSGLTKRSTRTSIHAPGDODR	
	KLGEDLENNEVNPEVSEFPIVLCILNSTPEYVVGIFCAIINGQIPAFSIIFSR	
	FTREDEPETKRONSMFSLVFLVGIISITFEFLQGFPGAGEIILKRLRYVFER	
	LROVSMFDPDKRNTTGALTTRLEADNOVGAGISRLATVIONANTGQIISLYG	
	WQTLULLAIYPIITAIAGVYEMKLSQALKDKKELAGCATEALEENRTVYSLR	
	EOKETVMQSLQVPIRNSLRKAHIFGVSEITQDAMVFSTAGCFRGATLVANEPNN	
	FOQDLVFAISVIFGAMAVGVSSPAPDYAKAKVSAAHVIMIEKSPILDSYSPGLK	
	NTELVNATFNEVNVNPTRPDIPIVLOGLSILEVKSGQTLALVSGSGKSTVYOL	
	YDPLAGSVLIDGKEIKHLNOMLRANHIGVSIILDCSIAENIAYGDSNRVSH	
	IMWAKENIHHIFETLPEKNTVRVKGIOQLSGGQKORIAIARALROPHILLDE	
	TSALDTSEKRYVQDALRKARGRRCIYIAHLSTIIONADLIVRONKVAHEGHQ	
	LAOKGIYFSMYSVQAGAKR"	

QY	560	tggaatttaaaatgttcttcttcaattatccatcaagaccatctatcaagattcttgaag	619
Db	1197	tggaatttcaaaaattgttcttcttcaatttcttccatcccttcttggaaaagaattgaattctttaaag	1256
QY	620	gtctaatctcgaattaaagctctcggagacagctgccttgctgcggtcttcaatggagacgt	679
Db	1257	gtctcaacttgaaaggttcagagttggcccaacactgtggccttggttggaacatgctgctgcc	1316
QY	680	ggaagaqtaagtaagtcacagctctctcagaaggttataigtatccgga tga tggctttatca	739
Db	1317	ggaagagcagcagaccgctgacgcagctgacagagcgcttatgaccaccacagatggcatgtct	1376
QY	740	tggttgatagaatgaatgaatcagaagctttaaagtctgcgcatatctgagaccatatggag	799
Db	1377	gtaattcattgacagacgaactatgagacattgaattgaatgaatgacattcttcggaaatttactgtg	1436
QY	800	tggttgatcagaagcctgtgtgtcttgcggaccacca tca gtaacaaatcaatgaatgagac	859
Db	1437	tggtagtgcagagcctctgtgtgttggccaccacgattactgtaaaaatttcgtatggcc	1496
QY	860	gagatgaatgtaactgta tgaagatgagagagcagcaagggagcaaatgctgatgt	919
Db	1497	ggcaaaattgtcaccaatgagatgagattgacaaaacctgttaaggaagcaaatgcttatgatt	1556
QY	920	ttatcatggaatttccctaataattaaatgaatttgtagtgagaaaagagctcaaatga	979
Db	1557	ttatctatggaactgaactaatatgaatttgaactctgtgttgagagagagggcccaactga	1616
QY	980	gtgagagagcaaaacagagagatcgcaattgtcgttgcttaagttcgaaaccccaagatc	1039
Db	1617	gtggtagcagaaacagagagatcgccattgtgcctggccctctggttgcgaaccgaatttc	1676
QY	1040	tgattttagatgagctacgctctgccttggaattcagaagaacagtcagctgttcaagctg	1099
Db	1677	ttctcgtgatgagcgcacagctcgtcgtgacactgaaatgtaagcagtgcttcagctgg	1736
QY	1100	caactgagaagagagcaagaagtcgagactacaatctgtgtagacacaccgacttcta	1159
Db	1737	ccttgatataaggccacaataaaaggccgagactacatttgatagtgctacatggttgctacag	1796
QY	1160	ttcgaagtcgagaatttgattgtgacctaaagaatgtaatgctgcggcggaagaaagagcac	1219
Db	1797	tttcgtaatgcgcaattgtcattgcgtggtttgatgatgagctgatttgcgaataaaggaaatc	1656
QY	1220	atgctgaactaaatgycanaaacgagctcatatatacttaacttgatgtaagcagata	1279
Db	1857	atgatgaactcaatgaagaaagaaaggcatttacttcaaaactgttcacaaatgcacaaagag	1916
QY	1280	aaaaagctgatga -acagatgtagtcaatgaacatatctctcgtgaagaagaagacaactca	1338
Db	1917	gaaatgaaaatttgattgaagaaatgccactgctgtaattccaanaatgtaaaatgtaagtccttgg	1976
QY	1339	cttcctcctgcaactctg tgaagagc-----atcaagtcagaactcaatgtaacagctg	1390
Db	1977	aaatgcttccaaaaagatttgcagggtccagtttaattaaaaaagaatcaactgcgacagagta	2036
QY	1391	aggaatcacaaccaatctaagaagataagctt-----	1422
Db	2037	taacatgcaccacacaaagccaaagacagaaagcttggtaacaaaagagacttgaatgagaa	2096
QY	1423	--ccctgaagctctcatataaaaatttaaagttaaacaagcctgaatgacgttttg	1480
Db	2097	taactccagtttcttcttgcagagattctgaagctgaactcaactgaatgagccttatttttg	2156
QY	1481	ttctcgggacatcgtgctctgcttcaaatgaagactgtlcatcaagtaatttccatcatc	1540
Db	2157	tggtttggtatatttttgctcatattataaagcagagccctgcgaaccagcatttttcataattat	2216
QY	1541	ttgcaaaaattacaactgtt--ggaataatgataaacaacacataaagcatgatg	1597
Db	2217	ttttcaagagatttataggagcttttaccggaatgatgactcttgaaacaaacaaagcagagata	2276

QY	1598	cagaattatcatgatatcttcgtcaatttcgttggttgatattgccttcagttatctta	1657
Db	2277	GTAACTATCTTTTCGATATCTTTCTCTAGCCCTTGGAATATATCTTTATATCAATTTTTC	2336
QY	1658	tcgaaggatattttacacggacagacagggaaattttaaacgtatgacttaagacacttg	1717
Db	2337	TCCAGGCTTACATTTTGGCAAAAGCTGGGAGATCTCTCACTAAGGGCTTGATATCATG	2396
QY	1718	cccttaaacgaatgtatatacagatatcttcgtgttttgaatgaagaaagaaacagcaag	1777
Db	2397	TTTTCAATTCATCTGTGAGACAGAGATGTGACGTGGTTTATGACCTTAAAAACCACTG	2456
QY	1778	gaagcttgaacaacaattttgccaatagatagacaacaattaaaggaacaacagttcca	1837
Db	2457	GAGCAATTACACACCAGGCTTGCCAAATGATGGCTCTCAAGTTAAAGGGCTTATAGGTTCA	2516
QY	1838	ggattggcgtcttaacaacaatgaacactacaatggagacttcacgtatcatcttcctta	1897
Db	2517	GGCTTGCCTGATTAACCCAGATATATAGCAAACTTGGCAAGGCACTTTATATCTTAA	2576
QY	1898	tatatgaatggagagatgacatctccatgattctgaatgatctccagtaactgcgtgaag	1957
Db	2577	TCTATGGTTGGCAATTAACCTTTTACTCTTTAGCAATTTGACCTCATATGCAATAGAG	2636
QY	1958	gaatgatgaacccgcgcaatgaatgcatttcgttccaaagaatatgaagaacactaacg	2017
Db	2637	GAGTGTGTTGAATGAATAAATGTTGCTGGAGACAACTGAAATATGAAGAAGCTAGAG	2696
QY	2018	aagcttgaagaatagacaacatgaagcctttgagagataatacgtactaagtgtcatatcaaa	2077
Db	2697	GAGCTGGGAATATTGCTATACGAAGCCATGCAAAATCTCGAAGCTGTGTCTTTGAACTC	2756
QY	2078	gggaaaaagccttcgagcaaatgtatagaagagatgcttcagaactacaacagaaatact	2137
Db	2757	GGGACCAAAATTTGAATATACATGATATCAACAGAGTTTGGCAATACCAACAGAAACTCTT	2816
QY	2138	cgaagaagaacacatatatggaagcgtgtatgacatgcagatgcctttatataatttg	2197
Db	2817	TGAGGAAGACACACTCTTCGGGGCTCATATTTCTATACCCAGCAAGATGATATTTTT	2876
QY	2198	ccatacgaacagggcttcgatattggaacatttaatacgaactgaacgaatgaccacag	2257
Db	2877	CCTATGCTGGCTGTTTCCGGTTTGCTGCTACTTGGTGCAATGATGATCAAGAACTTTC	2936
QY	2258	aaggcaatgtacatgatttttaactgaatgacatgacatgagcatatggccatcggaaaacgc	2317
Db	2937	ACGATGTTCTTTTGGTATATTCACACATCTGTCTTGGTGCATATGCGAGTGGGGGAGTCA	2996
QY	2318	tcgatttgctcccgtaatatccaaacgaacacggggcttcgcatcgttttgctctgt	2377
Db	2997	GTTCAATTTGCTCTGACTATGTCCAAAACCAAGTATACAGAGCCACGTCATCATGATCA	3056
QY	2378	tgaagaagaacaacaatatagacacgcgcgactcaagaagaaggaaagccacagacatgtg	2437
Db	3057	TTGAAAAAGCCCTGTGATTTACAGCTACAGGCCCTCAGCGCCCAAGCAATACTGTTGG	3116
QY	2438	aagggaattcagagtttcgagaagctccttcttcatccatgltgcgccagatgtttca	2497
Db	3117	AAGGAAATGTACATTTATATAGTGTGCTGTTCAACATATCCACTGCAGCAGACATCCCG	3176
QY	2498	tccttcggtgcttatccctcaatgataatgagaggaagaagaglaagatttggggagaa	2557
Db	3177	TGCTCTCAGGGGTGAGCTTCAGATGGAAGAGGCCAGACGCTGGCCCTCGTAGGTATCA	3236
QY	2558	ggagctcgttgggaaagacactcgttcaaatcttcagagagacttgaaccccgcgacag	2617
Db	3237	GTGGCTGTGGGAAGACCAATTTTACGCTCTAAGCGCTTCTATATGACCCCTTGGCTG	3296
QY	2618	gacaagatcgtttagatgtgtgagacaaagaatgaatgtacagatgctcgtctcc	2677
Db	3297	GTTCAAGTGCTAATTGATGGCAAGAGATTAAGCAACCTGATGTCAGTGGCTCCGAGAC	3356
QY	2678	aaatagcaatcgttccctcaagaagccgtgtccttcaactgcagcaatgtcgtgagaacatcg	2737

Db	3357	ACCTGGCATCTGCTCTCCAGAGACCCATCCCTGTTTGACATGCGAGCATTTGCCGGAACATCTG	3416
OY	2738	ccta tgg tga caa cag cgt g t g t g c c a t t a g a t g a t c a a g a a g c c g c a a t g c a g	2797
Db	3417	CCATTGGAGAACAAACGCCGGGTGTTACACATGTAAGATTTATGCAAGCCAGCGCAAGAGG	3476
OY	2798	caaatcattcttcttatttgaagctctccctgagaatatacaacacaaagltgactga	2857
Db	3477	CCAACTATACACCACTTCATCTGAGACACCTCCCTGAGAAATACACACACAGATGAGACACA	3536
OY	2858	aaggagacagactcttcggcgcccaaaacaagactagctatcttgcaaggctctcc	2917
Db	3537	ANGAACCCACGCTCTCTGGTGGCCAAAGAACGCGCATTCCTCCATACCTCGCGCTCTTGTTA	3596
OY	2918	aaaaacccaatatttatttgatgtaggcacttcagccctcgatataatgaactgga	2977
Db	3597	GACAGCCATATTTTGTGCTTTTGGATGAAGATCATCATCACCCTGTGATACAGAAATGTAA	3656
OY	2978	aggtggttcagcagbccccttgataaagccaagacggaagaacatgctagtgtgctactc	3037
Db	3657	AGGTTGTCCAGAGGCCCTGGACAAAGCCAGAAAGGCCGCGCACCTGATTTGGATCGGCC	3716
OY	3038	acagacctctgcaatttcagaacgcagatttgatagtggttctgcacaatggaagataa	3097
Db	3717	ACCCCTTGTCCACCATCCAAAGTCAGATTTAATTAATCTGTTGATGAGATGGCAAAAGCA	3776
OY	3098	aggacaagaacatcatcaagagctccctgagaatcgacaatatatttgaattgag	3157
Db	3777	AGGAGCATGGCGACACATCAACACACTGCTGGCCACGAAGACGATCTATTTTCCATGCGCA	3836
OY	3158	atgcacagctcag	3169
Db	3837	GTGTCCAGGCTG	3848

RESULT	4			
AXIOS082				
LOCUS	AXIOS082	4279 bp	DNA	30-Apr-2001
DEFINITION	Sequence	26	from Patent WO0123540.	
ACCESSION	AXIOS082			
VERSION	AXIOS082.1	GI:13921232		
KEYWORDS				
SOURCE	dog.			
ORGANISM	Canis familiaris			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
AUTHORS	1 (bases 1 to 4279) Stocker,P.J., Stelmel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.			
TITLE	P-glycoproteins and uses thereof			
JOURNAL	Patent: WO 0123540-A 26 05-Apr-2001;			
FEATURES	GENEST CORPORATION (US)			
source	Location/Qualifiers			
	1..4279			
	/organism="Canis familiaris"			
	/db_xref="taxon:9615"			
	17..3862			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC37723.1"			
	/db_xref="GI:13921233"			
	/translation="MDPEGGRKSAEKNFWMGKSKKKKKKKKKPKYSPFAMPRYSWLDRLVHLVGTMAIIHSGALPLMLLYFGKNTDSFANAGISRNKTFEYLINESTNNQOHEFNHDEEMTYTAAYYSIGIGAGVLVAAYIQVSPFCIAARQLILKQEFHIMQCEIGFEDWDVGEILRLFDQSGINEGIGKGFSPICAFEPFGFVIFGRWKLRTIVLLASIPVLGSAIIMKILISFEDKELAAKAGAAVEVLAIRVIARFGCKEELIERYKNLEAKGIGIKAKITANISIGAAFLITVASYALAAEYWGTSIVLSSEPTGOVUTFPESSAIGSIOASPIEAFANRGAAYEIFKLIIDNKPISIDSYSSKSGPNDINGNLEKRNHAFSPSRKQVYILKGLNLKVOGOTVALVGNSSGCKSTVOLMORLYDPDIDGMVCIIDGODITFIVNRHLREITGVASDPEVLFATTLNENIRYGRNWTMDIEAAVKEANADFTLKEPNKEDTLVGERGAQLSGGQKRIARVLYNPKLILLDEATSLDIEEASEVAOVAUOLDKARKGRTTIVAHSTRYANDVAFEDGVYIEKGNIDELMKKGIYI			

Query Match	35.6%	Score 1130:	DB 6:	Length 4279:
Best Local Similarity	62.4%	Pred. No. 2.7e-222:		
Matches 1879:	Conservative	0:	Mismatches 1085:	Indels 48: Gaps 5:
BASE COUNT	1296 a	833 c	1009 g	1141 t
ORIGIN				
QY	200	ttgtaagatcacagaatctccaagaatctgcaagaatcttgccataaagaagactatag	259	FKLVYMTQNGEITLENATGSKSESDALEKSPKDGSSSLIKRSTRSRSHAPQODR
DB	843	TTGAAAGGTACACAAAATTTAGAAGAGCTAAAGAAATTGGATTAAGAAAGCTATCA	902	KLGRKEDLNENPVPSWFLIKLNTSTWPFVVGIVFCALINGGLQAPESITFSRIIGT
QY	260	cttcaaaagtgcctcttggtgctgctactcttctatgaatgaaacctatgactgtct	319	KTREDDEPTKRONNSMFSVLEVLGISTEIFFPQGTTFGKAGELIRLRYFRSM
DB	903	CGGCCAACATTTCTATTGGTGCCCGCTTCTTATTGATCATATCATATGCTGTGCTT	962	LRODWSMEDDKRNTTGALTITFLANDAAQVGAISRLATVONIANLRKRGVFRSM
QY	320	ttgtgatggaacctcccttgatcttctaataygaaacctggatataccatcggaactgtc	379	MOLLLILAIIVPIAIGAVYEMKMLSGOALKDKELGACKIAEALLENRTVAVSLR
DB	963	TCGTGTATGGGACCTCTGTGGTCTCTCCAG-----TGAATATCTATTGGACAAGTAC	1016	QOKREYNAOSIOYPRYNSLEKATIRPVYSITQAMMYEYACGRCGATVLANEPFN
QY	380	ttgcgtttctttagtgaatccataagatcttgatcttgatggagcagcagctccact	439	FQDVLFSALVFAGMAVAVGVSSFPADISLEVKRGQTLVALIVSSSGCKSTVVOLEFR
DB	1017	TCACGTCTCTTTCTGTGATTAATTGGGGCTTTTAGTATTGGACAGCATGCCAACCA	1076	NTLGNVATFENEVFNYPTRPDI PVLOQGISLEVKRGQTLVALIVSSSGCKSTVVOLEFR
QY	440	ttgaaaccttcgcaatagcccgagagagctgcctctcatacttccaaagttatgataa	499	YDPLAGSVLIDGKEIKHLNOMWMLRGLGISQEBILFDCSIAENIADGDSRVSHBE
DB	1077	TTGAAGCATTTTCCAACGCAAGAGAGAGCGAGCTTATGAATCTTCAAGATTAATTGACATTA	1136	IVQAKENAIHHLETPLEKYNTRGKRGDLOSGCKORIAIAALVRPHILLDEDA
QY	500	aaccagaatagataacttctccaaagcttgatataaactgaaacctcatalagaagaacty	559	TSALDTSEKVVQDALDKARRGRCIVIAHRLSTIQNADLIVRNQKVEHGHQDL
DB	1137	AACCAAGCATTTACAGCTATTGGAAGAGTGGACATAAACCAATATATTAAAGGAAATT	1196	LAOKGIYFSMYSVQAGAKR"
QY	560	tggaaattaaagaatttctctcaatitccataagaccatctatacgaatctcgaag	619	
DB	1197	TGGAAATTCAAAATGTCTACCTTCAGCTTACCTTCTTCGAAAMGAACTTAAAGCTTAAAGG	1256	
QY	620	gtctcaatccagaatlaagctctgtagagacagctgcgcttgctgcgttccaaatggcagtg	679	
DB	1257	GTCTCAACCTGAAGTTTTCAGAGTGGGACAGACAGTGGCGCTGTTGGAAACGTGCTCGG	1316	
QY	680	ggaagaatgcgtagtccagcttctcgaaggttatalgaatccggaatgagctttatca	739	
DB	1317	GGAAAGAGCGACGCGGACGTGATGACAGAGGCTCTATAGACCCCAAGATGAGCATAGCTCT	1376	
QY	740	tggctgtagagagaatcacatcagaacttcaaatgtgcgcatatctgagaacctatitgag	799	
DB	1377	GTAATGATGACAGAGACATTAGGACCATTAAGTAAAGGCAATCTTGGGAAATTAATCTGTGG	1436	
QY	800	tggtagtagtaagaagctgttctgttcgagagaccacatcagtaacaatatcaatgagac	859	
DB	1437	TGGTAGTACGAGAGCCTGTGTGTTTGGCCACACACATAGCTGAAAACATTGCCCTATGGCC	1496	
QY	860	gagatgatatgcatgtaatgaagaatgtagagagacgaagaagcaaatgcgtatgata	919	
DB	1497	GGCAAAATGTCACACGTGATGAGATTGAGAAAGCTGTTTAAGGAACCAATGCTCTATGATT	1556	
QY	920	tataatgtagagcttccataaattataacttggcttgtagggaaaaaaggagctcaaatga	979	
DB	1557	TTATATCATGAATACCTAATTAATTTTGAACCTCTGTGTTGGAGAGAGAGGGGCCACGCTGA	1616	
QY	980	atgagagagagaaacgaatgcgaatgtgcctgtaccccttacttgcgaaccaccaagatctc	1039	

Db	1617	GGGTGGACAGAAACAGAGATGGCCATTGGCTCGGGCCCTGGATTGCAACCCCAAGATTC	1676
Qy	1040	tgattctagaatggagctacgctctcccttgattccagaagaacgaactgctltcaactg	1099
Db	1677	TTTCGCTGGATTGAGCGCAACGCTCAGCTCTGGACACATGMAAGTAAACACTGGTTGACGGTGG	1736
Qy	1100	caatgsgaagcgagcaaaagtcgagctacatacgtlgtlagaacaccgacttctacta	1159
Db	1737	CCCTGATTAAGGCCACAGAAAGCGCGGACATACCAATTGTGATCTCATGTTGTCTCAAG	1796
Qy	1160	ctcgaattgcagatttgattgttacccttaaggaatggaaatgctgycggaagaagagac	1219
Db	1797	TTTCGTAAATGCGAGTCAATTCATTTGCTGTTTTGATCATGAGCACTATTTGGAGAAAGAAATC	1856
Qy	1220	atgcttgacaatbgbcaaaacgagctctatattcaacttgaatgagctgaagcatatta	1279
Db	1857	ATGATGAACATCATCAAGAGAGAACGGCATTTACTTCAAACTTGTCAATTCGACGACMAAG	1916
Qy	1280	aaaaagctgahga-acagatlgagtlcaatlgacatacttactgaaagaagaaccaacta	1338
Db	1917	GAATGTAAATTTGAACTGTGAAAATATGCCACTGCTGAATCCAAAAGTGAATGATGCTTGG	1976
Qy	1339	cttccctctgcaactctgtgaaagac-----atcaagtcagactctatgcaagctg	1390
Db	1977	AAATGCTCCAAAAGATTCAGGGGTCCAATTATATAAAAAGAAAGATCACTCGAGAGCTA	2036
Qy	1391	aggatccaccacaactcctaagatagaactcitt-----1422	
Db	2037	TACATGCACCAACCAAGGCCACAGACAGAAAGCTTGTTCMAAAGACACTTGAATGGAATG	2096
Qy	1423	--cctgaagctcctctatlaaaaaattlaaagtltaaacaaagcctgaatgacctltg	1480
Db	2087	TACCTCCAGTTCCCTCTGTGGAGGATTCGAGCTGAAGCTGAACACTGAATGGCCTTATTTGG	2156
Qy	1481	cttcggggacatttggctctcgtctctaaatggaaactgttcatccagatattccatcatc	1540
Db	2157	TGGTTGGTAAATTTTGGCTATTATTAACGGAGGCGCTGCAACACATTTTTCATTAATAT	2216
Qy	1541	ctgcaaaaataatataccaactgttt---ggaataatgatataaacacactlaaagaactg	1597
Db	2217	TTTTCAGAGATTATATGGAATCTTTTACCAGATGAGAGATCCTTAACMAAACACAGAGATA	2276
Qy	1598	cagaaatlaltccatgatatactgcatacttgggtgtatattgcttctlccagtlatttca	1657
Db	2277	GTAACATGTTTTCGTATGTGTTTCTGATGCTTGGAATTAATTTCTTTATTAATCAATTTTCC	2336
Qy	1658	tgcaggaattatttcaatgcgcagagcaggggaaatttlaacgatlgatlaagaacatttg	1717
Db	2337	TCCAGGGCTTCACATTTTGGCAAAAGCTGGGAGATCTCTCACTAAGCGGCTTGGAATACATG	2396
Qy	1718	ccttcaagaagcatlctlatcagaatgatgtgcctgrrttatgaagaaggaaaacagacag	1777
Db	2397	TTTTTCAGATCTCAGTACGACAGAGATGTACGTGGTTGTATGACCTTAATAAACACCACTG	2456
Qy	1778	gagggcttgacaacatacttgacatagatataagcaaaaatlaaagagcaaacagtlcca	1837
Db	2457	GAGCATGTGACACACAGGCTTGCCAAAGATGAGGGGCTCAATTAAGGGGCTATAGTTCCA	2516
Qy	1838	ggattgagcgtcttaacaacaaatgcaactaacaatgaggacttcagtlatcatlcttca	1897
Db	2517	GGCTTGCTGTCATTCACCCACAAATATATGCAAAATCTTGGGACAGGCACTTATATATCCCTAA	2576
Qy	1898	tatbtgatbtgagatgagaccltccgatctcgtatctgagatattgctcagtlacttgcytgcag	1957
Db	2577	TCATATGCTGGCAATTAACCTTTTACTCTTACCAATTTTATCCCAATTCGCAATAGCAG	2636
Qy	1958	gaatgatataagccgcagcaatgactgagatttgcacaaagaataagcaagaactaagc	2017
Db	2637	GAGTGTGTTGAAATTAATAATTTGTCTGGACAAAGCACTGAAATATTAAGAAAGCTAGAG	2696
Qy	2018	atgcttgaaagatagcaactgaaagcttltggagaataatacgtlactaaglyfcatlaaaca	2077

Dh	2657	GAGTGGGAAGATTGCTACAGAAGCCATCGAANAACCTTCCGAACGTGTTCCTTACATC	2756
Qy	2078	gggaaaaagccttcgcgcgaatgtatgtaagagatgcttcgaactcaacaagaaatacct	2137
Dh	2757	GGGAGCAGAAATTGAAATACATGTATGACAGAGATTTCAGATGCATACAGAAGACTTT	2816
Qy	2138	cgaagaagcagcagatattatgtgaagctgtltagcaatcaagcattgaccttatatttgg	2197
Dh	2817	TGAGAGAAAGCACACATCTCCGGGGTCTCATTTTTCATATCCACAGCAATGATATTTT	2876
Qy	2198	cctatgaagcaggtttgattttggagaccttatattcaagcgcgcgaatgaccocag	2257
Dh	2877	CCATAGCTGGCGTGTTCOCGGTTTGAGTCTACTTCTGTGGCAATGACTCATGAACTTTC	2936
Qy	2258	aggagcattcatagatttttactgcgaattgtcaiaegagcctatgcccactcogaanaacgc	2317
Dh	2937	AGGATGTCTTTTGGTATCTCAGATATGTCTTTGGTGCCATGCGAGTGGGGAGGTCA	2996
Qy	2318	tcgltttggtccctgaaatattccaaagccaatcgggggcgtgcacatcgtttgacctt	2377
Dh	2997	GTTTCATTTCGCTCCGTACTATGCCCCAAAGCCAAAGTATCAGACGCCACGTCATCATGTCA	3056
Qy	2378	tggaaaaaagaaaccaaatatagacagccgagcgaaggaaggaagaaacgcagacaacatg	2437
Dh	3057	TTGAAAAAGCCCTCGATGTAGAGGTATACAGCCCTCACGGCCTCAAGCCAAATACGTTGG	3116
Qy	2438	aaggaaatttagagtttcgagaagctcctttcttcatccatgltccgcacgaatgtttca	2497
Dh	3117	AAGGAATGTGACATTTAATGAGCTGTGTTCACACTTCCACATCCACATCGACAGACATCCCG	3176
Qy	2498	tcctcgttgcttaccctccatgattatgagcgaagaaagacaagtatgcatltgtgggagca	2557
Dh	3177	TGCTTCACGGGGCTGAGCGCTCGAGTGTAAAGAAAGGCGCAACGCTGGCTGTAGATGACA	3236
Qy	2558	ggcgctgtggaaaaagagcacttcgttcaactctgcagagacattatgaccocgtgcaag	2617
Dh	3237	GTGGCTGTGGGAAGAGACAGTGTTCACCTCTTGAGACCGCTTCTATGACCCCTTGGCTG	3296
Qy	2618	gacaagtgaccttattgtagtgtgtgtagtgcaaaagaaattgaatgtaacgtgacctgtccc	2677
Dh	3297	GTTTCAGTGTCAATTTGATGAGGCAAGAAAGATTAACACCTGTAATGTCCAGTGGGTCGAGAC	3356
Qy	2678	aaataagcaatcgltccctcaagagcgtgtgcttccaactgcagacattgcttgagaacatg	2737
Dh	3357	ACCGGCGATCGTGTCTCAGAGACCCCATCTGTTGACTGCGAGCATTTGCCGAGACATTTG	3416
Qy	2738	cctatgtgtgcaaacagccgtgtgtgtgcattatgaaatcaaaagaagcccgacaatgacg	2797
Dh	3417	CCCTATGGAGCAACAGCCGGGTGTATTCACATGAAGCAATTTGTCCAGGCGACCAAGAGAG	3476
Qy	2798	caaatcacatcttccttatttgaagtgctccctcgagaataaacacacaagcttgagactga	2857
Dh	3477	CCAACATACACACTTCATCTGAGAACACTCTCCGTGABAATTACACACACAGATGAGAGACA	3536
Qy	2858	aaggagacagccttctgctgcggccagaaaacaagaactagctatgtgcaagggctctctcc	2917
Dh	3537	AAGGACCCACAGCTCTGTTGGTGGCCAGAAACACGCAATTGCAATAGCTCGCGCTCTTTGTA	3596
Qy	2918	aaaaaacccaaatttatgtttgttagtagagccattcgagccctcgataaagacatgagac	2977
Dh	3597	GACAGCCTCATATTTCCTTTTGGATGAAGCTACATCATGAGCTCTGTGATACAAATAAGTTTAA	3656
Qy	2978	aggtgtgttcagcaatgcccttgataaagccagagacgggaagagacatgcctagttgcatc	3037
Dh	3657	AGGTTGTCCAGAAAGCCCTTGACAAAGCCAGAAAGGCCCACTGCAATTGTATTCGCC	3716
Qy	3038	acagcctctctgcaattcagaaagcagatltttagtattgttctgcacaattggaagaataa	3097
Dh	3717	ACCCGTTGTCCACATCAACACACTCTGTGGCCACAAAGGCAATCATATTTTCCATGGTCA	3776
Qy	3098	aggacaagaagcactcatcaagagctccctgtagaataatcgagacatatatttaagttagtga	3157
Dh	3777	AGGAGCATGTGCACATCAACACACTCTGTGGCCACAAAGGCAATCATATTTTCCATGGTCA	3836





Db 2079 AAATGCTCCAAAAGATTCAGGGTCCAGTTTAAATAAAGAGATCACTGCAGAGATA 2138  
 1391 aggaatccaccocatctaaagataagtcct----- 1422  
 Db 2139 TACATGCAACCAAGGCGAAGACAGAAAGCTGGTACAAAAGAGACTTGAATGAGAAATG 2198  
 1423 --ccctgaagctcctctatlaaanaattlaaagltlaacaagcctgaatgaccttctgtg 1480  
 2199 TACCTCCAGTTTCCCTTCGAGAGATTCTGAAGCTGAACCTGAACCTGAATGAGGCTTATTTTG 2258  
 1481 tctctggagacatttgctctctgttctaagtgaactgttcaatccaglatlccatcatct 1540  
 2259 TGGTGGATATTTTGTCTATTATTAACCGAGGCTCAACACACATTTTCAATAATAT 2318  
 1541 ttgcaanaattlaacacatgltt---ggaataatgataaaccatlaaagctgtag 1597  
 2319 TTTCAAGATATATAGGATCTTACCCGAGATGAGGATCTGAAACAAACACAGATTA 2378  
 1598 cagaattatcatalcatalcgtcatcttggtgtctatctgtctgtcagtlattca 1657  
 2379 GTAAATGTTTCTGTATTGTTTCTAGTCCCTGGAATATTTCTTTATTAATTTTCC 2438  
 1658 tgcaggatatttaccgagcagagcagggaaatttlaacgagtgatlaagacattgg 1717  
 2439 TCCAGGCTTCACATTTGGCAAAAGCTGGGAGATCTCCTCACTAAGCGCTTCGATACATGG 2498  
 1718 ccttaagaacgctgtatatacagagatctgctgtgttgatgaaaggaacagacag 1777  
 2499 TTTTCAAGATCCATGCTGAGACAGAGATGTCAGCTGGTTGATGATCCCTAAAAACACCATG 2558  
 1778 gagctctgaacaacatattagcatalagacacaaattcaagagacagacagttcca 1837  
 2559 GAGCATTTGACAAACAGGCTTGCCATGATGCGCTCAAGTTAAAGGGCTATAGTTCCA 2618  
 1838 ggaattgagcttcaacacaaatgcactaacaatgagacattcagttatcatcttca 1897  
 2619 GCGTTGCTGTCAATTAACCCAGATATAGCAAACTTGGGACAGCATTTATATATCTTAA 2678  
 1898 tatatgagtgagagatgagacatcccgatctcgtgagtatgctcagttactgctgcgagac 1957  
 2679 TCTATGCTTGGCAATTAACACTTTTACTCTTAAGCAATGTGATCCCATTTGTAATAGCAG 2738  
 1958 gaattatgaaacccgacgaacatgactggaattgccaacaaagataagaagacactaac 2017  
 2739 GAGTGTGTAATGAATAATGTGTCTGCAACAGCACTGAATAAGTAAGAAAGCTAGAGAG 2798  
 2018 atgctggaagatagcaactgagacttggagaaatatacgtactatagtgtaactaa 2077  
 2799 GAGCTGGAGAAATGCTACAGAAAGCCATCGAAACCTCCGAACGTGTGTTCTTGGACTC 2858  
 2078 ggaagaaagcctcgagcaaatgtaagaagatgcttcagactcaacaacaaatact 2137  
 2859 GGGACACAGAGTTTGAATACATGATGACAGAGTTTGGAACTGATACATACAAACTCTT 2918  
 2138 cgaagaagaacacagatatttgaagactgtatgcatcgaagcattgcttataatttg 2197  
 2919 TGAGGAAAGACACATCTTGGGGTCTCATTTTCTATCAACCAAGCAATGATGATTTT 2978  
 2198 cctatgacgacgggttctgatttggacatttaactcaagctgagcaatgacccag 2257  
 2979 CCTATGCTGGCTGTTCCGGTGTGGCTACTGTTGGCAATAGATTCATCAACTTTC 3038  
 2258 agggacatgtcatagtttctactgcaattgtcatatgagagctatggtcactcgaaaaagc 2317  
 3039 AGGATGCTCTTTGGTATTTCTCAGCTATTTGTTTGTGCTTGGCCATGGCAGGGGCGATCA 3098  
 2318 tgcatttgcctcctaataattccaagaacaaatcggggctcgcgactggttgctgt 2377  
 3099 GTTCAATTGCTCTCTACATGCAAAAGCCAAAGTATCAGACAGCCACGTCATCATGATCA 3158  
 2378 tggaaagaaccaaataatagacagccgagcaagaaaggaagaaagcagacacatg 2437

Db 3159 TTGAAAAAGCCCTCTGATTTGACAGCTACAGCCCTCAGGCTCAAGCCAAATACGTTGG 3218  
 2438 aagggaatttgaagtttcagagaagctcctctctctatccatgctgcgccaagatgttca 2497  
 3219 AAGGAATGTGACATTTTATAGGTCTGTCTCAACTATTCACCTGACACAGACATCCCG 3278  
 2498 tctcctggtgttaccctcagatattgagcgaagaaagacagatgatttggggagca 2557  
 3279 TGCCTCAAGGGCTGAGCTCCAGGATGAAGAGGGCCAGACAGCTGCGCTCTAGTAGCA 3338  
 2558 ggcgctgtgggaaaaagcactctgttcaactctgcgagagactttagaccctgtcag 2617  
 3339 GTGCTGTGGGAAGACACAGCTGTTCAGCTTCAGACGCTGTCTATACCCCTTGGCTG 3398  
 2618 gacaatgctgttgaatgtgtagtgcataaagaattgaatgtaagtggtcgtcc 2677  
 3399 GTTCAGTCTTATTTATTTATGAGCAAAAGATTAAGACACTGAATGTCCAGTGGCTCCGAGCAC 3458  
 2678 aaatagcaatcgttctcctaagaagcctgtgtctctcaactgcaagattgtgagacatcg 2737  
 3459 ACCTGGCATGCTGTCTCAGAGGCCCATCTGTTTGACTGCAGCATTTGCCAGAAACATG 3518  
 2738 cctatgttgaacaacagccgtgtgtgtccatagatgagataaagaagccgcaatcgag 2797  
 3519 CTTATGAGAGACACAGCCGGGTCTGATCATGAAGATTTATCCAGCACCCAAAGAGG 3578  
 2798 caaatatcattcttctttagaagctccctctgagaaatatacaacacaaagttgactga 2857  
 3579 CCAACATACACCACTTTCATGAGACACTCCCTGGAATTAACACACAGATTAGAGACCA 3638  
 2858 aagagcagacacttcttgccgagcgaacaaagactagctatttgaagggctctcc 2917  
 3639 AAGGAACCCAGCTCTCTGCTGGCAGAAACAGCGCATTTGCCATAGCTGCGCTCTGTTTA 3698  
 2918 aaaaacccaataatttatttgttgatgagcgcaacttcaagccctcgataatgaacgtgaga 2977  
 3699 GACACCCATATATTTGTTTGTGATGAAGCTACATCAAGCTCTGATACAAAGTGTAA 3758  
 2978 aggtgttcagacatgctcctgtataagccaggaagcggagagacatgcttagtgtaactc 3037  
 3759 AGTGTCTCAAGAACCCCTGAGCAAAAGCCACAGAAAGCCGACCTGCAATTTGATCCGCC 3818  
 3038 acaggtctctgcaattcacaagcagatttgaatagtggttctgcacaaatggaagataa 3097  
 3819 ACCGCTGTCCACATCCAGATGAGATTAATAGTGTGTTACAGATGAGCAAAAGTCA 3878  
 3098 aggaacaaggaactcatlaaagctcctgagaagaatcgagacatatatttgaatgtaga 3157  
 3879 AGGACCATGGCACACATCAACAGCTGTGGCTCAGAAAGCATATTTTTCATGATCA 3938  
 3158 atgcagatcag 3169  
 3939 GTGTCCAGGCTG 3950

RESULT 6  
 AF045016 4317 bp mRNA MAM 07-FEB-1998  
 LOCUS Canis familiaris multidrug resistance p-glycoprotein (MDR1) mRNA,  
 DEFINITION complete cds.  
 ACCESSION AF045016  
 VERSION AF045016.1 GI:2852440  
 KEYWORDS  
 SOURCE dog.  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE 1 (bases 1 to 4317)  
 AUTHORS Puel.O., Lepage.J.F., Alvinerie.M., Gallier.P. and Pineau.T.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1998) Pharmacology, INRA, BP 3, 180 Chemin de  
 Tournefeuille, Toulouse Cedex 9 31931, France  
 FEATURES Location/Qualifiers





Dh 2387 TCAGAGGCTTCACATTGGCGAAAGCTGGGAGATCCTCATTAGCGGCTTCGATACATGG 2446  
Qy 1718 cctcaagaagccatgtatcatcagatatgtccgtggttgaatgaagaagaaacagacag 1777  
Dh 2447 TTTTCAGATCCATGCTGAGACAGATGTCAGCTGTTGATGACCCCTGAAACACACACTG 2506  
Qy 1778 gaagcttgaacaatatlaagcatalagacaaatcaagagagcaaggttccca 1837  
Dh 2507 GAGATTGACAAACAGGCTTCCATGATGGCGCTCAAGTTAAAGGGGCTATAGTTCCA 2566  
Qy 1838 ggaatgagcgttcaacaacaataatgaactaactaaggaacttccgttcccttca 1897  
Dh 2567 GGCTTGGCTGCTATTACCCAGATATAGCAAACTTGGGACAGGCACTTATATCTTAA 2626  
Qy 1898 tataatgagtagagatgaacttccgttccgttgaatgtgcccagtaacttccgttgaag 1957  
Dh 2627 TCTATGCTGTCATTAACATCTTTACTTACCTTATGCAATTTGTCATCCATCTTTCATAGAG 2686  
Qy 1958 gaatgaatgaacccgacgaactgaactgttggccaacaagaataagcaagaacttaagc 2017  
Dh 2687 GAGTTGTTGAATGAATAATGTTGCTGGACAAAGCACTGAAGATTAAGAAAGAGTAGAAG 2746  
Qy 2018 atgtctgaagaataatgaactgaacttgaagataatgaactatagttcattaca 2077  
Dh 2747 GAGCTGGGAAGATTGCTACAGAACCACTGCAAACTTCCGAACCTGTTCTTGTGACTC 2806  
Qy 2078 ggaagaagaagccttcgagaagaatgatgaagaagatgcttcaagcacaagaagaataact 2137  
Dh 2807 GGGAGCAGAACTTTGAATACATGATGACAGAGTTTGCAAGTACATACAGAAACTTT 2866  
Qy 2138 cgaagaagaacagatatltgaagcgttltatgaactcaagcatttataatttg 2197  
Dh 2867 TGAGAAAGACACACATCTTCGGGCTCATTTTCTATACACCAGCATGATGATTTT 2926  
Qy 2198 cctatgaagcaggggttccgttgaacttgaacttcaatcaagcgtgagcaatgacccag 2257  
Dh 2927 CCTATGCGCGCTGTTTCGGTGTGCTTACTTGTGCGCAAACTGTAAGTAACCTTC 2986  
Qy 2258 agggcattgtcatagtttctactcaatgtgatgatgaactgaactgacatcggaagaagc 2317  
Dh 2987 AGGATGTTCTTTGATATTCACAGCTATTTGTTGTCATGACAGTGGGCGACGTCA 3046  
Qy 2318 tcaatttgcctcctgaatattccaaagaacacgagggcgtgacatcgttgccttct 2377  
Dh 3047 GTTCATTGCTCTGACTATGCTCAAGCAAGATATCAGACCCACGTCATCATGATCA 3106  
Qy 2378 tgaagaagaaccaaataatgaacgacgacgacgaagaaggaagaaagcagacacatgtg 2437  
Dh 3107 TTGAAAAAGCCCTCTGATTTGACAGCTACAGCCCTCAGGCGCTCAAGCAATATAGTTGG 3166  
Qy 2438 aaggaatltagagtttcgaagaagtccttcttctatccatgctgcagagatgttttca 2497  
Dh 3167 AAGGAATGTACATTTATAGAGTCTGCTTCACTATCCACATGACACATCCCG 3226  
Qy 2498 tccctcgttgccttccctcagatltgaagcagaagaacacagatgttggggaagca 2557  
Dh 3227 TGCTCCAGGGCTGAGCTCTCAGAGTGAAGAGGCGCACAGCTGGCCTCGTAGTAGCA 3286  
Qy 2558 ggcgcttggggaagaagccttcttcaacttgcagagaacttataccccgtgcaag 2617  
Dh 3287 GTGGCTGGGGAAGAGCAAGTTGTCAGCTCTAGAGAGCGGCTTATATACCCCTTGCTG 3346  
Qy 2618 gacaagtccttltgatgtgtgaatgaagaagaatgaatgaatgaatgttccgttcc 2677  
Dh 3347 GTTCAGTCTTATTTGATGAGCAAGATTAAGACACTTAATGTCAGTGGCTCCAGAGAC 3406  
Qy 2678 aaatgaacatgcttccctcaagaagccttgccttcaactgaagcagatgttgaagaactg 2737  
Dh 3407 AACTGGGAGATGCTGCTCAGAGGCCATCCTGTTGATCGACGATTCGCGAAGCAATTG 3466  
Qy 2738 cctatgtgtgaacaagcgttltgtgtccattagatgaatcaagaagcagcaatgcaag 2797  
Dh 3467 CCTATGGAGACAGACGCGGCTGATCATCATGAAGATTAATGACGACGACGCAAGGAGG 3526

Qy 2798 caaatccatcttcttattgaagtcctccctgagaataacacacacaaatgtgactga 2857  
Dh 3527 CCACATACACACACTTTCATGAGACACTCCCTGGAATATCAACACAGATGAGACACA 3586  
Qy 2858 aaggaacagcttcttgcgcgacagaacaagaactagctatttgaaggtcttcc 2917  
Dh 3587 AAGGAACCCAGCTCTGCTGTCAGAAACAGCCCATTTGCCATGCTGCGGCTCTTGTTA 3646  
Qy 2918 aaaacccaatttatttltgtatgaaggcacttcaagccttgcgtatgaagatga 2977  
Dh 3647 GACAGCTCATATTTTGTGTTGATGAGCTACATCAGCTGATGATCAGAAAGTAAA 3706  
Qy 2978 aggtgttcagcatgccccttgataaagccaggaaggaagacatgcttgaactc 3037  
Dh 3707 AGTTGTCCAAAGAACCCCTGACAAAGCAGAGAAAGCCGACCTGATGTGATCCGCC 3766  
Qy 3038 acaagctctctgcaatcagaagcagatttgaatgtgttgcacaaatggaagataa 3097  
Dh 3767 ACCCTTGTCCACCATCCAGATGATTAAGTGTGCTTTCAGAAATGGCAAACTCA 3826  
Qy 3098 aggaacaaggaactcatcaagagctcctgagaataatcgaacatatatttaagtga 3157  
Dh 3827 AGGAGCATGCGCACATCAACAGCTGCTGCTCAGAAAGCATATTTTCCATGATCA 3886  
Qy 3158 atgcacagtcag 3169  
Dh 3887 GTGTCCAGGCTG 3898

RESULT 7  
AX105059  
LOCUS AX105059 4317 bp DNA PAT 30-APR-2001  
DEFINITION Sequence 3 from Patent WO0123540.  
ACCESSION AX105059  
VERSION AX105059.1 GI:13921211  
KEYWORDS  
SOURCE  
dog.  
ORGANISM  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 4317)  
AUTHORS  
Stocker, P.J., Stelmel-Crespi, D.T., Crespi, C.L., Reif, T.C. and  
Patton, C.J.  
TITLE  
P-glycoproteins and uses thereof  
JOURNLS  
Patent: WO 0123540-A 3 05-Apr-2001;  
GENTEST CORPORATION (US)  
FEATURES  
source  
Location/Qualifiers  
1..4317  
/organism="Canis familiaris"  
/db\_xref="taxon:9615"  
70..3912  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC37719.1"  
/db\_xref="GI:13921212"  
/translation="MDPEGGRKSAEKNFMKMKSKKEKKPTVSTFAMFRYSNM  
LDRLYMVGTFMAALIHGAALPLMLIVGNNMDSFANAGISRNKTFPIINESTNNMO  
HFINHDEEMTYTAYYSIGAGVLAAYIQVSPKLAGSROLIKIKOEFHIMOE  
IGMFDVHDELNRLTLDVSKINBEGDVGFPSIATFTFGTIVGFPPKRLIY  
ILAI5PLVLSAAIWAIISSFTDKELAYAKGAAVEELAIRVIAFGGCKELE  
RYNKNLEAKRIGIKKAIITANISIGAFLIIVASIALAEWYGSLVSYTGQVLT  
VPSVLIGAFSISQASPIAFANRGAAYEIRIIDNKSISYSKSGHPNIGN  
LEFRNHFYSRSEKVKIILKGLNMGSGTVAIVNAGSGGCTTQLMORLDPDNG  
MVCIDGDDITINVRHLREITTVYSGEPULPATTIAENITRGRENMTMDIEKAVEA  
NAYDFIMKLPNKEDTTLVGERGARISGGKORIAIARALVNRPIILDEFTSALDES  
EAVVOVALDRNKRRTTIVIAHRLSTVRNADVLAGEDDGVIVKGNHDEMLKRGIFY  
KLVTMOTRGNIELENATGSEKSESDLEAPDSGSLIKRSTRSHIAPGODGR  
LGRKEDENVPYSFMRILKINSTENPVGIFCAIINGQAPAFSIIFFSLIGF  
TRDDEPTKRONSNMPSVILVIGTISFTIFFGCTEGGAGELIKRLRMYVRBML  
RQDSWFDDEKNTTGALTTLANDAAOVKAITSRLAVITQNTANLGTGIIISLIYGM  
QTLTLIAIVPIIAGVNVEMKLSGALKDKRELKAGIATIEALENFTVLSRE  
QKEHYMYAOSLIQVYRNSLRKAIIFGVFSITQAMMYFSVAGCFRCGATIVANEFNF





OY	920	ttaccatgagagcttccctcaataaatttaacatctggttagggaaaaagagctcaatga	979
Db	1959	TTATCATGAAACCTCCCATTAATTTTGACCCCTGGTTGGAAGAGAGGGCCGCTTGA	2018
OY	980	gtggaggccaanaacagagatcgcaatctgcgtgcgtctagtttcgaaaccccaagatc	1039
Db	2019	GTGGTGGCCAGAACCAAGATGCCATTGGACGTCCTGGTTGGCAACCCCAAGATCC	2078
OY	1040	tgattttagatgagctacgtctgcacctgcatccagaaaagccaagtcgtctcaactg	1099
Db	2079	TCCTGCTGGAATGAGGCCACGCTCAGCCTTGGACACAGAAAGCAGACGATTCAGTGG	2138
OY	1100	cactgsgaagagcgagcaaaagtcgcgactacaacgtggttagacacccgacttctacta	1159
Db	2139	CTCTGGATTAAGGCCACAGAAAAGTCTGGACACACCTTGTGTACTCATCTATGTTCTTACAG	2198
OY	1160	ctcgaatctgcagatttgatgtgtgacccttaagaatgaaatgctgcggaagaagaagac	1219
Db	2199	TTTCGTAATGCTGACGTATCGTCTGTTTCGATATGGAATCAATTGTGGAGAAAGAAATC	2258
OY	1220	atgcttgacaatctgccaanaacgagctctatatctacgttctgactgtaacagagatla	1279
Db	2259	ATGATGAACCATCAATGAAGAAGAAAGCATTTACTTCAAACTTGTCAAAATGCAGACAGAG	2318
OY	1280	aaaaagctgtagaacagatgtagtcaatgacalatctca-----ctga	1322
Db	2319	GAATGTGAAGTTTGAAATTCAGACGTCGATGAATCCAAAGTGAATTCATGCTGCTTGG	2378
OY	1323	aagaagaaccaactctctctgcacctgagagagacalca-----	1366
Db	2379	AAATGTCTTCAAAATGATTCAGATCCAGTCCTTAATAAGAAAAGATCAACTCGTAGAGTG	2438
OY	1367	-----agtcagacttcatgacaagcgtgagaatcccaaccatctaaagagtaatc	1420
Db	2439	TCCGTGATTCACAAGCCCCAGACAGAAAGCTTAGTACCAAAAGAGCTCTCGATGAATA	2438
OY	1421	ttctcgtagctctctctataaaaaattlaaagltlaaacaagcctgtaigtcccttctg	1480
Db	2499	TACCTTCAGTTCCCTTTTGAGAGATTAATGAACTTAATTAATGAAATGGCCTTATTTTG	2558
OY	1481	ttctggggacattggctctgctcttaaaaggaaactgtcatccagatatttccatcatct	1540
Db	2559	TTGTGTGGTGAATTTTGCCATTATTAATGAGAGCGCTCGAACACACATTTGGCAAAATAT	2618
OY	1541	ctgcaaaaatatataccaatgttctggaaaataatgaataaacacatlaag---catlgtg	1597
Db	2619	TTTCAAAAGATTATAGGGGTTTTTACAGAAATGATATCTCTGAACAAAACAGCAGATTA	2678
OY	1598	cagaatattatccatgatalatctgcataatttgggtgttatttgccttctgcagttattca	1657
Db	2679	GTAACCTGTTTTTACATATGTGTTTTCTPAGCCCTTGGAATTAATTTCTTTATTAATTTTCC	2738
OY	1658	ttcgagggaattatttlaagcgcaagacaggggaaatttlaacgatgagatlaagcaacttgg	1717
Db	2739	TTTCAGGGTTTTACATTTTGGCAAAAGCTGAGAGATCTCTCACAAAGCGCTCCCATACATGG	2798
OY	1718	cctcaaaagccatgtatatacagagatatctgcgtgttgaatgaaaggsaaaacagcaag	1777
Db	2799	TTTTTCCGATCCATCTCTAGACAGAGATGTAGTTGGTTTATATACCCCTAATAAACACCACTG	2838
OY	1778	gagagcttgcacacaaatlaagccatagatalagcacaaaaatlaaagaggsaaaggttcca	1837
Db	2859	GAGCATGTGACTACCAAGCGCTGCCAAATGATGCTCTCAAGTTAAAGGGGCTATAGTGTCCA	2918
OY	1858	ggatttgcgtcttaacacaaatgcaactaacatgtagacttcaagtlatcatcttcctta	1897
Db	2919	GGCTTGTGTATTAATCCCAAAATATGCAAAATCTTGCGACAGGAATTAATTATATCTTCCA	2978
OY	1898	tatatgagtggagaatgacatctcgtatcttgagtatgtctcagtlacttgcgcgtgaag	1957
Db	2979	TCTATGTGTGGCAACTAACCTGTACTCTTTCAGCAATTTTACCCATCATTTGATGATAGCG	3038
OY	1958	gaatgattgaaacccgcagcaatgactggtgttgcacaanaagataaagcaagaaactaagc	2017

[illegible]

RESULT	9
LOCUS	HUMMDR1
DEFINITION	HUMMDR1 4666 bp mRNA PRI 03-DEC-1999
ACCESSION	M14758
VERSION	M14758.1 GI:1874668
SOURCE	P-glycoprotein; drug resistance protein; transport protein. Human drug-resistant cell line KB-C2.5 cDNA to mRNA, clones lambda-HDR[10, 5, 104].
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 4646) Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B. Internal duplication and homology with bacterial transport proteins in the mdrl (P-glycoprotein) gene from multidrug-resistant human cells
TITLE	Cell 47 (3), 381-389 (1986)
JOURNAL	Cell 47 (3), 381-389 (1986)
MEDLINE	87028230
REFERENCE	2 (sites) Ueda,K., Clark,D.P., Chen,C.J., Roninson,I.B., Gottesman,M.M. and Pastan,I. The human multidrug resistance (mdr1) gene. cDNA cloning and transcription initiation
TITLE	J Biol. Chem. 262 (2), 505-508 (1987)
JOURNAL	J Biol. Chem. 262 (2), 505-508 (1987)
MEDLINE	87109132
REFERENCE	3 (bases 971 to 985; 3095 to 3109) Kloka,N., Tsubota,J., Kakehi,Y., Komano,T., Gottesman,M.M., Pastan,I. and Ueda,K. P-glycoprotein gene (MDR1) cDNA from human adrenal: normal P-glycoprotein carries Gly185 with an altered pattern of multidrug resistance
AUTHORS	Biochem. Biophys. Res. Commun. 162 (1), 224-231 (1989)
TITLE	Biochem. Biophys. Res. Commun. 162 (1), 224-231 (1989)
JOURNAL	Biochem. Biophys. Res. Commun. 162 (1), 224-231 (1989)
MEDLINE	89322246
COMMENT	[2] sites. Draft entry and computer-readable sequence [1] kindly submitted by I.B.Roninson, 13-AUG-1987. The sequence shown is of a cDNA clone initiating at a minor upstream transcription initiation site and containing the major site of transcription initiation. Location/Qualifiers 1. 4646 /organism="Homo sapiens" /db_xref="taxon:9606" /map="7q21" 1. 4646 /gene="PGY1" 382 /gene="PGY1" /note="g in [1]; a in [2]" 425. 4267 /gene="PGY1" /note="P-glycoprotein" /codon_start=1 /protein_id="AA59575.1" /db_xref="GI:307180" /translation="MDLEGDNGGAKKKNFKLNKSEKDKKKPTVSPMSRYSNWLDTLVGVGLAIIHGAGLPLMLIVGEGWTDIFPANGNEDLMSNTNSDINDYGFENLEEDMPYAAIYSGICAGVLYAAYIQVSPICLAAGRIHRIROFLAIRORFEGFVHDGELNRLTIDVSKINEIVIGDKIQMFOSMATEFTGIVGRWKLTLYVLAISPVGLISAAMVAKILISTFTDELLEAVAGAVAEVLLAIFTVLAFGQKLELR
FEATURES	
SOURCE	
gene	
variation	
CDS	

variation	964	/gene="Pgy1"	/note="c in [1]; c in [2]"	978..979	/gene="Pgy1"	/note="t in [1]; ga in [2] val->Gly"	1660	/gene="Pgy1"	/note="c in [1]; t in [2]"	2065	/gene="Pgy1"	/note="c in [1]; t in [2]"	3101	/gene="Pgy1"	/note="g in [1]; t in [2] Ala->Ser"	3859	/gene="Pgy1"	/note="c in [1]; t in [2]"	4460	/gene="Pgy1"	/note="a in [1]; g in [2]"	BASE COUNT	1371 a	892 c	1129 g	1254 t	ORIGIN	154 bp upstream of Aval site; chromosome 7q21.1.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
variation																					Query Match	35.4%;	Score 1124;	DB 9;	Length 4646;	Best Local Similarity	62.1%;	Prod. No. 4.6e-221;	Matches 1869;	Conservative	0;	Mismatches 1095;	Indels	48;	Gaps	4;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
variation																					QY	200	ttgtgaagtatacaagaatctccaagaatgtcaagaatctttgtgcataaagaagactatag	259	YNNKLEARRKRIKKAKATTANISGIAFLLIYASALAFWVGTTLVLSGSEYLSGOVLLT PFSVLICAEFSVGGQSPSEIEFAPNRRGAAYEIPKTIIDKRPISDYSKSGHKRPNDIKGLP EPRVAFHESVPSRKAEVKTLKGLNLRKVSGQTVFALVNGSGCKSTTVOLMOQLDPTBGM VSVQGDQIRTIINWFLEKEIIGVVSQEPVLPAITIAENIRRGREWTVIMDELEKAEKN AYPEIMKLPKRPDLVGERGAQLSGGQKORIAIALRVNRPKILLDEEASALDTESE AMVOALDKARKGRTTIVIAHRLSTVANAIVSGFDGVIEKGNHDLKEKGIYK LIMOTAGNEVELEENADESKSEITDALEMSINSGSRSLIKRRSRBRSSQODRL STRKALDESTIPVPSFMRIMKLNLTMEPYFPVGVCAIINGLOPAPAFITFSKILGPT RIDDPEFKRONSLPESLFLAALGIIISITFPVLOGFTFGAKGELIKRLRIWFSMLR QDVSWFDPKNTGALTTFLRLANDAAQKGAIGSLAVITANTIANLKGIIISFLYGO LLTLLAILVPIIALAGVEMKMLSGQALFKRKLEEGKGLATEIENFRVSVLTOBO KEFMYASLOSLVOPRNSLRRAHIFIGITFEFTQAMMYSVAGCFRGAIVLAHKLMSE DVLIVFAVVGAMAVGQVSPADYKAKKISAOTIIMIEKPLIDSTYEGTEGMPMT LEGAVTGEVYVETPRPDIPLVQGLSELEKPOILFDOSIENBAYGNSRVYSOEELV PLAKVLLDGEIKRLNVQWRRAHLGILSVKPELFDOSIENBAYGNSRVYSOEELV RAAKEANIHAFTESLPKRYSTVKGDCTOSGGOKORIAIALRVORPHLLIDEATS ALDSEKERVQOELDKAREGRTCTIVIAHRLSTIONADLIIVFONGRVKEHGTHQOLLA OKGIFYSMSVQAOAKTRQ"																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	

D	1599	TGCAATTTCAGAAATGTTTCACCTTCAGTTACCCATCTTCGAAAGAAAGATTAAGATCTTGAAG	1658
Q	620	gtctgaatctccgaatlaaagtcctgagagacagtcgccttgcgttcgtctcaatgagcagtg	679
D	1659	GCCGAACCTGGAAGGTGCAGAGTGGGCGACAGACGGTGCCCTGGTTGGAAACAGTGGCTGTG	1718
Q	680	ggaaagtaecggtgaagtcacagctctcgcagaagttatagatcccgatgaatgagcttatac	739
D	1719	GGAAGACACACACAGTCCAGCTGATGTCAGAGGCTCTATGACCCCAAGAGGGATGTGTCA	1778
Q	740	tgatgtaaggaaatgacatcaagagctttaaagtcgcgcattatccagaccataatgag	799
D	1779	GTCGTGATGACACAGATATTAGACCATTAATGTAAGCTTTCTACGGGAAATCATTTGGTG	1838
Q	800	tgatttaaccaagagcctgttcttgcgggacacacatcagtaacaataatcaagatagac	859
D	1839	TGGTGAATCAGGAACCGTATTTGTTGGCACACAGATAGCTGAATAATTTGGCTATGGCC	1898
Q	860	gagatgaatgactgatgaaagagatgagagagcagcaaggagaacaatgcgatgat	919
D	1899	GTCGAAATGTCACACAGATGATGAGATGAGAAGAGCTGTCAAGGAAACCAATCCATAGCT	1958
Q	920	ttatcaaggagtttccctaataatlaacttctgtgtagggaaaaagagctcaatga	979
D	1959	TTATCATGAAAGCGCTCTAATTAATTTGACACCCCTGGTTGGAAGAGAGGGGCCAGTTTGA	2018
Q	980	gtcggaggcagaaacagagatcgcaattgcgcgtgcctttagtccaaccacaagattc	1039
D	2019	GTCGTGGGCGAAAGCAGAGATGCGCATTTGCACGTGCCCTGTGTCACACCCAGATTC	2078
Q	1040	tgattttagatgagcgtacgctctgccttgcattcagaagcaagtcagctgttcaagctg	1099
D	2079	TCCGTGCTGATGAGGCCACGCTGACGCTTGAGACACAGAAAGCAAGCAGTGGTTGAGGTG	2138
Q	1100	cactggagaaagcgagcaaaagtcggaactcaatcgtgtgtagcacaccgacttctacta	1159
D	2139	CTCTGGATTAAGGCCAGAAAGGCGGACCCACCATTTGTATACCTCATGTTTGTCTACAG	2198
Q	1160	ttcgaatgtagaatttgattgtgacccctaagaatgtagaattgtgcggagaagaagac	1219
D	2199	TTGCTAATGCTGACGTATCCGTGCTGTTTGATGATGATGACTCTTTGTGAGAAAGGAATC	2258
Q	1220	atgcgtacaatgycaaaaacgagtcctatattactcacttgcataagagatatta	1279
D	2259	ATGATGACATCAGAAAGAAAGGACATTTACTTCAAACTTGTCACAACTGCAGACAGCAG	2318
Q	1280	aaaaagctgtagaacagatgtagtcaatgaatatctca-----ctga	1322
D	2319	GAAATGAGTTTGATTAAGAAAGCGAGCTGATGAATCCAAAGTAAATGATGCTCTGG	2378
Q	1323	aagaagaaccaactcctctcctgcgaactctgtaagagcatca-----	1366
D	2379	AAATGCTTAAATGATTCAAGATCCAGTCTTAATTAAGAAAAAGATCAACTGCTAGGAGTG	2438
Q	1367	-----agtcgaactcattgacaagcgctcgagaaatccaccacatcaagaagataag	1420
D	2439	TCCGTGATCACAAAGCCCAAGACAGAAAGCTTAGTACCAAGAAAGGCTCGATGAAGTA	2498
Q	1421	ttctcgaagctctctataaaaaatttaagttaaacaacgcctgaatgaccttctgtg	1480
D	2499	TACCTCCAGTTTCTTTTGGAGGATTAATCAAGCTAAAATTACTGAAGGCTTATTTTG	2558
Q	1481	ttcggggacaatggctctcgtcttaaatggagacgttcatccagatatttccatcact	1540
D	2559	TTGTTGGTGTATTTTGTGCCATTAATAAGGAGGCTGCAACACACATTTGCATAAATAAT	2618
Q	1541	ttcgaanaattataacaatgltttggaataatgatataaaccacatcaag---cagtg	1597
D	2619	TTTCAAAAGATTAAAGGGGTTTTACAAAGAAATTGATGATGCTCAAAACAAAGACAGATA	2678
Q	1598	cagaatattatccagatattccgtcaattgtgggttataattgcttctgcagtatcca	1657
D	2679	GTAACCTGTTTACACATGTTTCTTACGCCCTGGAAATTTTCTTTTATACATTTTTC	2738

QY	1658	tgagaggaattatttcagcagagcagcggaattttaaagctlbgatlaagacacttgg	1717
Db	2739	TTTCAGGGTTTACCAATTTGGCCAAAGCTGGGAGACATCCTCCCAAGCGGCTCCGATTCATATG	2798
QY	1718	ccttcaaaagccatgtatatacagatatatgcgcgtgttttbtatgaaaaagaaaaacagcacag	1777
Db	2739	TTTTTCGATTCATCGCTCAGACAGCATGTGAGTTGGTTGATGACCTATAAAACACCACTG	2858
QY	1778	gaagcttgcacaacaatatatagccatagatatagcacaaatctaaaggagcaacagtttcca	1837
Db	2859	GAGCTTATCACTACCGAGGCTCGCCAAATGATGCTGCATCAATTTAAAGGGCTATATAGTTTCCA	2918
QY	1838	ggatttgcgccttcaacacaaaatgcacaactcaactatgggaacttcagttacatcttcccta	1897
Db	2919	GGCTTGCTGTATTTACCCAGAAATATAGCAAACTTGGAGCAGGAATATATATCTCTTCA	2978
QY	1898	tatatlgatlgggagatlgacatctccgatctlgagiatlgtccagtlacttgcgtlgacag	1957
Db	2979	TCTATGTGTGGCACTAACACTGTACTCTGTAGCAATTTGACCATTCATTTGCAATGTGAC	3038
QY	1958	gaatgatlgaaacccgcagcaaatgactgtgatltgccacaacaagataagcaagaacttaagc	2017
Db	3039	GAGTGTGTGAATGAAATGTTGTCTGGCAAGCAACTGAACATATAAGAAAGCACTGAAG	3098
QY	2018	atgtcgggaagaatagcaacttgaagcttbgagaatalaagtaactatagttgtcatlaaca	2077
Db	3099	GTGCTGGGAGATTCGGTACTGAAAGCAATATGAAAACTTCGGAACCGTGTCTTTGACATC	3158
QY	2078	gggaaaaagccttcagcaacaatgtatgaagaagatgcttcagacccaacaacagaataact	2137
Db	3159	AGGACACAGAGTTTGAACATATGTATGTGCACAGTTTGACAGGTACCATACAAACCTCTT	3218
QY	2138	cgaagaagaagacagttatlggaagctgttatgcacttgcagtcagtcacttatatatatttg	2197
Db	3219	TGAGGAAAGCACACACTTTTGGAAATTAATTTTCTTCACCCAGGCAATATGATGATTTT	3278
QY	2198	ccatagcagcagaggttccgattctggagccatlaaataaagctggacgaatgaacccag	2257
Db	3279	CCTATGCTGAGATGTTTCCGGTGTGGAGCTACTTGGTGACATATAACTCATAGAGCTTGG	3338
QY	2258	agggcatagttcaatgttcttactgcattgtcatatgagagctatgagcactatgcacacga	2317
Db	3339	AGGATGCTCTGTATATATTTTACGCTGTGTCTTGTGCTCCATATGGCGTGGGCAAGTCA	3398
QY	2318	tcgatttgcctcctgaataatccaaagccaaatcggggctgcgcgactcgtttgccttgt	2377
Db	3399	GTTCATTTGCTCTCTGACATATGCCAAAGCCAAATATACAGACGCCCATCATCATATGATCA	3458
QY	2378	tggaaaagaacaacaatatagacagccgcagttcaagaagaagaaaaagccaagacacatgtg	2437
Db	3459	TTGAAAAAACCCCTTTGATTTGACAGTACAGCACGGAAGGCGCTTAATGCCGACACTTTGG	3518
QY	2438	aaggaatatlagatgttctgagaagatcctcttcttcaatccatcagtcgcccagatgtttca	2497
Db	3519	AAGGAATATGTCATTTTGTTGTAAGTTGTATTTCAATATATCCACCCGACCCGACATCCGAG	3578
QY	2498	tcctccgttgcttataccctccagtatltgacgagaagaagacagtagacatltgtltggagaga	2557
Db	3579	TGCTTCAGGAGACTGAGCCGCGAGAGTGGAAGAAAGGGCCAGACGGCTGTGTGGGCGACGA	3638
QY	2558	gcgcgttggggaagaagcaactctgttcaactcttcgacagagactttatgaccccgltcaag	2617
Db	3639	GTGGCTGTGGAAGAGCACAGTGTCCACCTCTTGAGGCGGTTTCAAGCACTCTTGGCAG	3698
QY	2618	gacaaagtgtgttatagtgtgtcgtgacgaaaaaagaattgaaatgtacagtgatccgtctcc	2677
Db	3699	GGAAAGTGTGCTTATGTATGSCAAGAATAAAGACGATGAATGTTCAGTGTGCTCCGAGCAC	3758
QY	2678	aaatagcaatcgttccccaagagcctgtgtcttcaacttcagatgcagcatgtctgagaacatcg	2737
Db	3759	ACCTGGGCACTCTGTGCCCGAGAGCCCATCCTTTTACATGCAAGCAATTTGCTGTGAACATTG	3818



QY	2738	cctaagtggaacaaacagccgtgtgtgtccattagatagatcaagaagccgcaaatgacg	2797
Db	3819	ccatattggaacaaacacccgggtgtgtacacgaaagatcgtgaaggccagcaaaagacg	3878
QY	2798	caaatatcatcttcttatttgaagtcctccctggaataatcaacacaaagttggnactg	2857
Db	3879	ccaaacttaccatgcttcttaccatgactcactgcttatttatttaccactttaaagtagaaca	3938
QY	2858	aaggagacacagcttctcgcgcgcaaaaacaaagactagctatctgcaaggtcctctcc	2917
Db	3939	aaggaaactcagctcctctggtggccacaaacaaacgcatgttgcattagctcggtccctgtta	3998
QY	2918	aaaaacccaaatttatttattgttgatgtaggccaactcagccctcgataatgacatgaga	2977
Db	3999	gacagcctcatattttgtcttttgatgaaagccacgctacgctctgacaaagtgtaaa	4058
QY	2978	aagtggttgaagtcgccttgataaagccagagccggaagagatgcctagtgtaactc	3037
Db	4059	aggttctccaagaaagcccttgacaaagccagaaagagccgacctgtgattgtgttgcct	4118
QY	3038	aaaggtcctctgaattcgaagacagcagatctgaatgtagtctgcgaataagaaagaa	3097
Db	4119	accgctgtccacacacatccgaatgacagacttaattagtggtttcagaatggcagactca	4178
QY	3098	aggaacaaagaaactcatcaagagctcctcgaagaatcgagacatatatttcaagttagta	3157
Db	4179	acgagcattggcagcagctacagacactgctgtgcacagaaagcattatttttcaatgtctca	4238
QY	3158	atgcacagtcag 3169	
Db	4239	gtgtccagagctg 4250	
RESULT	10		
LOCUS	108557	4669 bp	PAT 02-DEC-1994
DEFINITION	Sequence 3 from Patent WO 8705943.		
ACCESSION	108557		
VERSION	108557.1	GI:588735	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 4669)		
AUTHORS	Rominson, I.B., Pastan, I.H. and Gottesman, M.M.		
TITLE	COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS		
JOURNAL	Patent: WO 8705943-A 3 08-OCT-1987;		
FEATURES	Location/Qualifiers		
Source	1..4669		
BASE COUNT	1394 a 892 c 1129 g 1254 t		
ORIGIN			
Query Match	35.4%;	Score 1124;	DB 6; Length 4669;
Best Local Similarity	62.1%;	Pred. No. 4.6e-221;	
Matches 1869;	Conservative 0;	Mismatches 1095;	Indels 48; Gaps 4;
QY	200	ttgtaaggtatatacagaatctcaagaatgcaagaatttgatcataaaagactatag	259
Db	1245	tttgaaaggatcaaaacaaatatttgaaagaaacttaaaagaaatttggaataaagaaagctattta	1304
QY	260	cttcaaaagtgtctctgtgtgtgttactctttaaagtaagaaacttagagactgtc	319
Db	1305	caagcctaattttctatagagtgctgttcttctgctgactatgacatcttattgtctgtgacct	1364
QY	320	tttgatatagaacctctctgtatcttcaatgagagaacctgatatacatcggagctgtc	379
Db	1365	ttctggtatgggacacacttgggtctctcagggca-----atatattgattggacaaatgac	1418
QY	380	tttgctttctctttagtgaatccatagcagttatitgcatgtgagcagcagctccctact	439
Db	1419	tcactgtatttcttttctgtattatttgaggcctttttagtggtagacagcattctccaaaca	1478
QY	440	ttgaaccttcgcaatagcccgagagctgtccttccatatttccaggtatgtataaga	499
Db	1479	tttgaaagatttgcaaaatgcagagagacgcttttgaaattcttcgaattattgtatata	1538
QY	500	aaaccagtatagataactttccacagcttgatatataaacctgaaatccatagaagaaactg	559
Db	1539	agccaaagtattgacaaagctatttcgaagagtgaggcaaaaccaaagaaatttttgaaggaattt	1598
QY	560	tggaaattaaagaatgtcttccatattatccatcaagaacccatcaagaatctcgaag	619
Db	1599	tggatttcagaaattgttctacttccagtttaccattctcgaanaagatttgaattcttgaagc	1658
QY	620	gtctgaactcgaatlaaagctctggaagagacagtcgccttgctgtgtcattcattgacgtg	679
Db	1659	gcctgaacctgaagtgtagagtgaggcagacagtgagccctggttggaanaacagtgctgtg	1718
QY	680	ggaagatgaagtgatccagcttctcgaaggttatatgataccgagatgagtgttattca	739
Db	1719	ggaagagacacaaacagtcagctgtatgacagagccttaagaccacaaagagggagtgca	1778
QY	740	tgtgtgagtgaaagaaacacagcagctttaaagtgtgcgcatatcgaagacatattgag	799
Db	1779	gtgttgatgacacagatatttgacacatttgaagtttcttctacaggaattcatttgctg	1833
QY	800	tgtttagtcaagagacgtgttctgtctcggagacacacatcaagtaacataatcaagttagac	859
Db	1839	tgtgtgagtcaggaacacctgtattgttgccacacagatagctgaanaacattgcgttagcc	1898
QY	860	gagatgatgtcactcgatgaaagatgagagagcagcagaaggaagcaatgcgtatgtc	919
Db	1899	gtgaaatgtgcacatggatgagatttggaagagcgtcacaagaccatgcttatgact	1958
QY	920	ttatcatgagttccctcaataatttaacatttgtagtggggaagagagctcaaatgga	979
Db	1959	ttatcatgaacactgccttatataatttgacacctgtgttgagagagagggccacgttga	2018
QY	980	gtgagagcagaacaaacagagatcgcgaattgtcgtgtcctttagtctggaaccccaagattc	1035
Db	2019	gtgtgtggcgaacaaacagagatgcgcgaatttgacacgctgtgttgcaaccccaagattc	2078
QY	1040	tgaatttagatgaggtacgtctgcctctggaatcgaagaacgaatcagctgttcaagctg	1095
Db	2079	ttctgtctgagatgagggccacgctacacgttgacacagaaagacgaactgttgcagctgg	2138
QY	1100	cacgtggaagagcagcagaaggtgagactaaatgtgtgtgagacacagacttactata	1155
Db	2139	cttctgattgaagcgaagaaagagtcgacacacattgtatgtatgctatgtgtgttgcag	2198
QY	1160	ttcgaagtgcagatttgaattgaccccaagagatggaatgcctgtgcgagagaagagac	1219
Db	2199	tttctgtatgacgtacgtcatcctgtgttgcattgatgagatcatttgtgagaagaaatgc	2258
QY	1220	atgctgaaactaatgycaaacagaggtctatatattcaactgtgtgatacagagatatta	1279
Db	2259	atgattgaactcatgaaacagaaagcattttcttcaaaactgtgtcacaatgacagacgacg	2318
QY	1280	aaaaagctgataagacagatgtagtaatagaatatctca-----ctcga	1322
Db	2319	gaaatggaagttgataaaatgacgtgataatgtccaaanaagtaattgattgctcttgg	2378
QY	1323	aagaagaagcaactcactctctcgtacctctgtgaagagcatca-----	1366
Db	2379	aaatgtcttcaaatgatttcaagatccaattcattatgaagaaaaaagatcaaacactgcctagagatg	2438
QY	1367	-----agtcagactctatgacaagcgttgagaaatccacccaactaaagataaagtc	1420
Db	2439	ttccgtgattacaaagcccaagcagaaagcttagtaccaaagagcctctgattgaagta	2496
QY	1421	ttcctgaagtcctctatlaaaaaatttaaagttaaacaaagcctgaaatgaccttltgtg	1488
Db	2499	ttacctccagtttcccttttggagagattatgaagctttaaattttaaactaaagcccttatttgg	2556





OY	260	ctctaaagtcgtctgtgtgtcgtgtactctcttatgaatgaagaaacctatgacttgc	319
Db	2697	cagccaaatattttcatatgagtgctgtcttccctgcctgatcatcattatctctgtgcct	2756
OY	320	tttgtatggaacctctcttatctcttaatgggaacctgatatccactgggaacttgc	379
Db	2757	ttctgattgggaccaccttctgtctctctcagggga-----aatattctattggacaaagac	2810
OY	380	tctgtcttctctttagtgatccatagcaagtatgcatttgatggacaagctccctca	439
Db	2811	tcacgtatttcttcttgcattattatggggccttttgattgttgacagacattccaaagca	2870
OY	440	ttgaaacctcgcgaatagcccgagagctgccttccataatttccaagttatataaga	499
Db	2871	tttgaaagcattttgcaaaatgacaagagagacagcttatgcataattcttcaagaattgtatata	2930
OY	500	aaccagatagataaacttttccacagctgagatataaactgaetccatagaaagaa	559
Db	2931	agccaaagtattgacagctatttccaaagtgggccacaaaccaaataattttaaaggaatt	2990
OY	560	tggaaattaaanaatglttcttccaattatccaaagcaacatccatcaagaatctcgaag	619
Db	2991	tggaaattcagaaatgtctcacttgcatttaccacattcgaaaaanaagtttaagattcttgaag	3050
OY	620	gtctgaattccagaattaacttcgagagagaagctgccttctgtgtctcctaatgga	679
Db	3051	gcctgaaccttgaaagctcacaagtgggcagacggctgcctctgttggaaacatgctctgtg	3110
OY	680	ggaagagatcaggttagctcagctcttcgcagaagttatcatgatccgagatgatgtcttca	739
Db	3111	ggaagagacaaacagctccacactgatgcagagccttatgaccccaacagggagatgtca	3170
OY	740	tgtgtgagatgagaatagacataagagctttaagtgtcgcgatatccgaacctatlgag	799
Db	3171	gtgtgtgatggacagagatttatgagacacataaatttaaggttttgcgggaaatcatgtgtg	3230
OY	800	tgtgttagtcaagagccctgttcttgcggagacccatccagtaacatatcaagtatgac	859
Db	3231	tgtgtgatgcagagAACGTGATTGTTGGCCACCGATACTGATAAACATTGGCTATGGCC	3290
OY	860	gagatgatgtgactgtatgtaaagagatgagagagcagaaggaagcaaatgcgatgatt	919
Db	3291	gtgaaatattgtccaccatggatgagatttgagaaacctctcagaagaccatgcttatgact	3350
OY	920	cttcaatggagtttcccaataaatttaatacatgttgtagggaaaagagactcaaatga	979
Db	3351	tttatcatgaaaactcctcatataaattttgacacctgtgttggaagagagggcccaagtga	3410
OY	980	gttgagagggcagaacacagaagatcgcgaattgtcgttctttagtttcgaaaccccaagattc	1039
Db	3411	gttgagggcagaaacacagagatgcattgcattgcacgtctcctgttgcaaaccccaagattc	3470
OY	1040	tgatattgaatgaagagctgtctgcaccttgattccagaagaagcgaactgttgaagt	1099
Db	3471	ttcttcttgatgagggccacacgtcagcctttggacacaaagaaacgaacagatggcttgggg	3530
OY	1100	cacttgaagaagcgagcaaaagctcgcgaactcaaatcgtgtgtgaacaacacgacttctacta	1159
Db	3531	cttctgataaagccagcaaaaaggtcggaccacattgtgtatgactcatgttggctacag	3590
OY	1160	tttgaagtgcgaatttgaattgtgaaccttaagaatggaatgtctgcgagagaaggaagcac	1219
Db	3591	tttctgaatgtcgaagctatgcgtgctgttgcagatntgaactttgtggaagaaagaaatc	3650
OY	1220	atgcgaactaatgccaacacgagctcatatattcactttgtatgtccaaggaattata	1279
Db	3651	atgataaactcattgaagaaagaaagcattttacttcaaaccttttccacatgcagacagacg	3710
OY	1280	aaaaagctgatgaaacagatgagtgatcaatgacatatctca-----ctga	1322
Db	3711	gaaatgaagttgaatttggaaatagcagctgatatgaatcctcaaaagtaaaattgattgcttgg	3770

QY	1333	aagaagaaccaactcaactlccctctgcaactctgtgaaagacatca-----	1336
Db	3771	AAATGCTCTCAAAAGATTCAGAGTCCATTAATAAAAAAGATCACTGTAGACGTG	3830
QY	1367	-----agtcagacttcattgcaagaagctgagaaatccacccaatctaaagataatc	1420
Db	3831	TCGGTGGATCCAAAGCCCAAGACAGAAAGCTTAGTACCAAAGGCTCTGGATGAAACTA	3890
QY	1421	ttccctgaagctccctctctctaaattttaaagttaaacaagaacctgaatgaccttltg	1480
Db	3891	TACCTCCAGTTTCCTTTTGGAGGATTATGAACTAAATTTAACTGAATAGCCCTATT	3950
QY	1481	ttctgggagcaatlygctctctgtctctaaatggaactgttcaatccagatattccatcatc	1540
Db	3951	TTGTGTGGGTATTTTGTGCCATTATAAATGAGAGCCCTGGACACACACTTTGGCAAAATAT	4010
QY	1541	ttgcaaaattataccaatctgtttggaaataatgatatbaaacaactaaag---catgtg	1597
Db	4011	TTTCAAGATTATATGGGGTTTTCACAGATTATATATCTGTAAACCAACACAGATTA	4070
QY	1598	cagaattatcccatgatatctgcacatlttgggtgttcatlttgcctgtcagttatcca	1657
Db	4071	GTAACTGCTTTTCACTATGTGTTCTAGCCCTTGAGATTATTTCTTTTATACATTTTTC	4130
QY	1658	ltgcaggatattttacgcgcaagacaggggaaattttaaagatlgatatgaacatltg	1717
Db	4131	TTTCAGGTTTACATTTTGGCAAAAGCTGAGATCCTCCACCAAGGGCTCGATACATG	4190
QY	1718	cccttaagaagcatgttatatccagatatctccgtgtttgataagaaagaaacagcaag	1777
Db	4191	TTTTCCGATCCATCTCAGACAGAGTGTGTTGTATATACCTTAATAAACACCACTG	4250
QY	1778	gaagctttacaacatatlttgcaagatatgaacaataatcaagaggaacaagttcca	1837
Db	4251	GAGCTTACACTACCAAGGCTGCCAATGATGCTGCTCAATTTAAAGGGCTATAGTTTCA	4310
QY	1838	ggattggcgtcttaacaacaatlygcaactaaacatgggacttcaagttatcatlccctta	1897
Db	4311	GGCTTGTGTAATTACCCAGATATAGCAAAATCTTGGACAGGAATAATTATATCTTCA	4370
QY	1898	tatatgattggagatgacatctctgaattctgatatgtctcaagttactgacgtcggaag	1957
Db	4371	TCTATGTTGGCAACTAACACTGTACTCTTACCAATTTGTATGCCATCATTTGCATAATGAG	4430
QY	1958	gaatgattgaaccccgcaagcaatlygcaatttgcacaacaataagcaagaactaagc	2017
Db	4431	GAGTGTGTTGAAATGAAATATGTGTCGTGACACACACTGAAAGATGAAGAAAGACTAGAG	4490
QY	2018	atgctggaagaatagcaactgaaagctlttgagaataatacgtactataglytcatbaaca	2077
Db	4491	GTGCTGGGAATGCTGCTAGTGAAGCAATAGAAACTTCGGAACCGTGTCTTGTGACTC	4550
QY	2078	gggaaaaagccttcgagcaaatgatatgaagaagtgctcaagactcaacacgaaatcact	2137
Db	4551	AGGACACAAAGTTTGAACATATGATGCTCAGATTTGGAGGTATCCATATACAAACTCTT	4610
QY	2138	cgaagaagaagcagatatttggaaagctgttatcatctcaagccaatgoccttataattg	2197
Db	4611	TGAGGAAAGCAACACTCTTGGAAATTCATTTCTTCACCCAGCAAGATGATGATTTTT	4670
QY	2198	ccatgacgacaggglttcgaatttggagcctatttaatccaagctggaacgaatgaccag	2257
Db	4671	CCTATGCTGAGATGTTTCCGCTTGGAGACCTACTTGTGTGACATATAACTCATGACCTTGG	4730
QY	2258	agggacatgttcataagtttttaactgacatgcatataggaactatgacctggaataaagc	2317
Db	4731	AGGATGTTCTGTAGTATTTTTCACACTGTGTTCTTTGGTCCATGCGCGGGGCAAGTCA	4790
QY	2318	tcgatttggcgcctgaatatattccaagaacaaatccggggctcgacatcgttctgtctgt	2377
Db	4791	GTTCAATTGCTCTGACTATGCCAAAGCCAAAATATCAGACGCCCATCATCATGATATCA	4855
QY	2378	tggaaaaaacaacatatagcaagcccgacgtcaagaagggaaaaagccagacacatgt	2437

```

Db 4851 TTGAAAAAACCCTTGTGATTGACAGCTACAGCAGGAGGCCCTAATGCCGACACATG 4910
OY 2438 aagggaatttagagtttcgagaagtccttcttctatccatgctgcccagatgtttca 2497
Db 4911 AAGGAATTCACATTTGGTGTGAATGTATTCACATTCACCCACCGGACATCCAG 4970
OY 2498 tctccgtggtatcccccagatattgagcgaagaagcagtagatgttgaggaga 2557
Db 4971 TGTCTTAGGGACTGAGCTCGAGGTACAGAGGCCAGCCTGCTGCTGCTGCTGCTG 5030
OY 2558 gctgctgtggaagaacactctctgttcaactctcagagacttltgaacccgtgcaag 2617
Db 5031 GTGGCTGTGGAGAGACAGAGTGTCCAGCTCCTGAGAGGGTCTCAGACCCCTTGGCAG 5090
OY 2618 gacaagtctgtttagtggtgtgtagtcaaaagaattgaatgtagtgcgttcc 2677
Db 5091 GGAAAGTGTGCTGTGATGTGCAAGAAATTAAGGAGCTGAATGTTCAGTGTGCTCGAGCAC 5150
OY 2678 aaatagcaatcgttctcgaagagcctgtgttcaactgcagcagcattgtctgagaacatcg 2737
Db 5151 ACCTGGGCACTCGTGTCCAGAGAGCCATCTGTTTACTGACAGCATTTGCTGAGACATTTG 5210
OY 2738 cctatgtgacaacagcgtgtgtgtgccaatagatgagatcaaaagccgcaaatgcaag 2797
Db 5211 CCTATGAGACACAGCAGCGGTGTGTACAGGAGAGATCTGTGAGGCGCAGAAAGAGAG 5270
OY 2798 caaatatcattcttcttattgaaagtctcctcgtagaataatacaacaagaattggaactga 2857
Db 5271 CCACATCATATGCTTTCATGAGTCACTGCTTAATTAATATGACACTTAAGTGAAGAGACA 5330
OY 2858 aaggaacagactctctcgtgcgagcaaaagaagactagactatgcaagggcctctcc 2917
Db 5331 AAGGAACATGAGCTCTGTGTGTGCGCCAGAAACAGCATTTGCCATAGCTGCTGCTTGTGA 5390
OY 2918 aaaaaccacaatttattgtttagatgtagagccactcgaacccctcgaataatgacagtga 2977
Db 5391 GACAGCCTCATATTTTGTGTTTGTGATGAGACCCAGTCAGCTGTGATGACGAAAGTGA 5450
OY 2978 aggtgttcaagcatgccccttgataaagcagaagcaggaagacaatgctagtgtaactc 3037
Db 5451 AGTTTGTCCAGAAAGACCTTGAGCAAGAACAGCAAGAGCCGACCTCATTTGATTTGCT 5510
OY 3038 acagagctcctgcaattcgaacgaacagacttgaatgtgttctgcacaatggaagaata 3097
Db 5511 ACCGCTGTCCACATCCAGAAATGACAGACTTAATAGTGTGTTTCAGAAATGGCAGAGTCA 5570
OY 3098 aggaacaagaactcaagaagctccttgaagaatcgacaatatatttaagtttagtga 3157
Db 5571 AGGAGCATGGCAGCATGACAGAGCTGCTGGACAGAAAGCATCTATTTTTCATAGTGTCA 5630
OY 3158 atgcacagtcag 3169
Db 5631 GTGTCCAGGCTG 5642

```

RESULT 12  
LOCUS AR028672 9318 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 6 from patent US 5858744.  
ACCESSION AR028672  
VERSION AR028672.1 GI:5940645  
KEYWORDS  
SOURCE

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 9318)

AUTHORS Baun, C., Stocking-Harbers, C. and Osterlag, W.

TITLE Retrieval vector hybrids and the use thereof for gene transfer

JOURNAL Patent: US 5858744-A 6 12-JAN-1999;  
FEATURES Location/Qualifiers  
1..9318  
Source /organism="unknown"

BASE COUNT 2377 a 2217 c 2364 g 2358 t 2 others  
ORIGIN

Query Match 35.4%; Score 1124; DB 6; Length 9318;  
Best Local Similarity 62.1%; Pred. No. 4, 6e-221;  
Matches 1869; Conservative 0; Mismatches 1095; Indels 48; Gaps 4;

```

OY 200 ttgtaagttataacgaatctcaagaatgcaagaagattttgcatataaagaagactatag 259
Db 2596 TTGAAAGGTAACAACAAATTTAGAAAGCTTAAGAAATGGGATTAAGAAAGCTATTA 2655
OY 260 ctcaaaagtgtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 319
Db 2656 CAGCCAAATTTCTATAGGTGCTGCTTCTGCTGATCTATGATCTTATGCTTGTGCTT 2715
OY 320 ttgtgataggaactcctctgattcttcaatgagaacctgatalacatcoggaactgttc 379
Db 2716 TGTGATGAGGACACCTTGCTGCTCTCAAGGGA-----ATATCTATTTGACAGATAC 2769
OY 380 ttgtgttcttctttagtgaatccatagcagttatgtcatgttgagcagcagtcctccact 439
Db 2770 TCACGTATTCCTTCTGATTAATTTAGGGGCTTTTGTGAGTGTGAGCAGGATCTCAAGA 2829
OY 440 ttgaaccttcgcaatlagcccgagagcgtgaccttcatatattccaggtatattgaaga 499
Db 2830 TTGAACATTTTGAATTCGACAGAGAGCAGCTTATGAATCTTCAAGATTAATTTGATTA 2889
OY 500 aaccagatagataaactttccacagcttgatataaaccgtgaatccatagagaagactg 559
Db 2890 AGCCAAGTATTTGACACCTATTTCGAAAGTGGGCACAAACACGATTAATTAAGGAAATTT 2949
OY 560 ttgaatttaaaatgttcttcttcaatattccatcaagaacatcatcaagaattcttaag 619
Db 2950 TGGAAATTCAGAAATGTTCACCTTCATGCCATCTGAAAGAAAGTGAATTCATTAAGG 3009
OY 620 gtctgaatcctcagaatlaagctcgtgagagagacgtcgacctgtgtgtgtgtgtgtgtgt 679
Db 3010 GCCTGAACCTGAAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 3069
OY 680 ggaagatcagtgtagtccagctctcgtgagaggtatataatgataatccgagatgactatca 739
Db 3070 GGAAGAGCACAAACAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 3129
OY 740 ttgttgatgaatgacatcagaagctttaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 799
Db 3130 GTGTTGATGACAGGATATTAGGACCATTAATTAAGTGTTCATGCGGAAATCATTTGGTG 3189
OY 800 ttgtttagtcaagagcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 859
Db 3190 TGGTGAATGACGAACTGTATTTTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 3249
OY 860 ggaatgattgactgaatgaagaatgagaagacgaagaagaagaatgactgactgact 919
Db 3250 GTGAAATGTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3309
OY 920 ttatcatgagtttctctataaatttaataatctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 979
Db 3310 TTATCATGAAATGCTCTCATTAATTTTGTGACACCTGTGAGAGAGAGAGGAGGAGGAG 3369
OY 980 gtgagagcagaagaagagatgcgaattgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1039
Db 3370 GTGGTGGGCAAGAGAGAGATGCGCATTTGCAAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 3429
OY 1040 tgaatttagatgagtgacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1099
Db 3430 TCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3489
OY 1100 caatgagaagagagagagagagagagagagagagagagagagagagagagagagagagag 1159
Db 3490 CTCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3549
OY 1160 ttggaagtgcagattttagtgcacctaaagaatggaatgtcgtgtgtgtgtgtgtgtgtgt 1219

```

Db	3550	TTGCTAATGCTGACGCATACGGTGGTTTCGATGATGAGTCAATTTGTGAGAAAGGAATC	3609
Oy	1220	atctgaacatcagcgcaaaacgagcctatatctatctctctgcatcgaagata	1279
Db	3610	ATGATGATCACTGAAAGAAAGAAAGGCACTTTACTTCAACTGTGTCAATGAGACACAG	3669
Oy	1280	aaaaagctgatacagatgagcaatgacatalctta-----ctga	1322
Db	3670	GAATGAAGTTGAAATAGAAAAATGCGCTGATGATCCAAAAGTGAATGATGCTTGG	3729
Oy	1323	aagaagaccaatctctccctctgacctctgtgaaggatca-----	1366
Db	3730	AAATGCTTCAATATGATTCAGATCCAGCTATATAGAAAAAGATCACTCTAGAGAGTG	3789
Oy	1367	-----agtcagactcctatgacaagcctgagaaaccaccaactcaagataagtaagtc	1420
Db	3790	TCCGTGATACACAGCCCAAGACAGAAACCTTAGTACCAAGAGGCTGTGATGAAGTA	3849
Oy	1421	ttctcgaagctctctattaaaatltaagltaaacaagcctgaatgyccttctgtg	1480
Db	3850	TACCTCAGTTTCCTTTTGAGAGATATATGACCTAAATTAATTAAGGCTTATTTTG	3909
Oy	1481	ttctgaggacattgagctctgtctctaatayggaactgltccagtatltccatctc	1540
Db	3910	TTGTGGTGATTTTGTGGCCATTATTAAGGAGGCTGCAACACCAATTCATTAATAT	3969
Oy	1541	ttgcaaaatataacaaatglttggaataatgataaacaacatatag---catgatg	1597
Db	3970	TTTCMAAGATATATAGGGGTTTTCAGAAATGATGATCCTTAACAAACAGACAGATAT	4029
Oy	1598	cagaattatltccatgatalatctgcatattltggltgltatltgcttltgcagtattca	1657
Db	4030	GTACTTGTTTTACATATGTGTCTTCTAGCCCTTGGAATATTTCTTTATATACATTTTCC	4089
Oy	1658	ttcgaggatattlttacggcagagcgaggggaaatttaagatgatatgaacacttgg	1717
Db	4090	TTTCAGGGTTTACATTTTGGCAAGCGTGAGAGTCTCCACCAAGGGCTCGATACATAGG	4149
Oy	1718	ccttcaagcgatltlatcagagatatgcccgtgttgaagaaggaaacagacag	1777
Db	4150	TTTTCCGATTCATCTCAGACAGAGATGTGAGTTTGTTATATACCTTAAMAAACACACTG	4209
Oy	1778	gagcgctgacacaacatlttagcctatagatalagcacaaatlcaagagcaaacaggttca	1837
Db	4210	GAGCATTTGACTACACAGGCTCGCCAAATGATCTCTCAAGTTAAAGGGCTATAGTTCCA	4269
Oy	1838	ggattgcgctctaacaacaatgcaactaacaatgagacttcagttcatltccctta	1897
Db	4270	GCGTTGCTGTAATTAACCCAGATATATAGCAAAATTTGGGACAGGAATATATATCTTCCA	4329
Oy	1898	tatatgagltggagatgacatctctgatatctctgagtlttcagactgtccgagag	1957
Db	4330	TCTATGTTGGCAACTTAACCTGTACTCTTAGCAATGTAGCCATCATTTGATGAATAGAG	4389
Oy	1958	gaatgattgaacccgcagcaatgcatgatttgcacaagaataagcaagaactaagc	2017
Db	4390	GAGTTGTTGAAATGAAATGTGTGCTGTGCAACACACTGAAAGATTAAGAAAGAACTAGAG	4449
Oy	2018	atcgtggaagaatagcaactgaagcctltggagaatalaagctactaagtgltcatacaa	2077
Db	4450	GTCGTGGGAAATGCTACTACCAATATAGAAAACTCCGAACCGTTGTTCTTGGACTC	4509
Oy	2078	gggaaaaagcctcgagcaaatgtatagaagatgcttcagactcaacagaaataact	2137
Db	4510	AGGACACAAAGTTTGAACTATGTATCTCAGAGTTTGACAGGTATCATTCACAAACCTTT	4569
Oy	2138	cgaagaagcgagataltatggaagcgtgtatgacatcaagcctgtcctatatatttg	2197
Db	4570	TGAGGAAGCAACACTCTTGTGGAATTCATTTTCTTCACCCAGGCAATGATGTATTTT	4627
Oy	2198	cctatcgagcagggcttcgatattgagacctatlttaactcaagctgagaaatgacccag	2257

Db	4630	CCATAGCTGAGTATGTTCCGGATTTCGAGCCCTACTGTGTGGCACAATAAATCATATGACGTTTG	4689
Qy	2258	aagggcatgttcacagtttacttactgcgaatgcatalcgtagcgtatgcccacatggaanaagc	2317
Db	4690	AGGATGTTCTGTTAGTATTTTTCAGCGTGTGTTTGGTGCCATGCGCGTGGGACATGCA	4749
Qy	2318	tgcgtttgctcctgcgaatattcccaagccaaatcgggggctctgcgcacatcgtlttggccttgc	2377
Db	4750	GTTTATTATTGGCTCTCGACTATATGCCCCAAAGGCCCAAAATATACAGACGCCACATCATATCATGATCA	4809
Qy	2378	tggaaagaagaaaccaaatatagagacgcgagtcgaagaagggaagaaagccagacatg	2437
Db	4810	TTGAAAAAACCCTTTTGATTGACAGCTTACAGACGGAAGGCCCTAATGCCAACACATTTGG	4869
Qy	2438	aaggaatttagagtttcgagaagtcctcttctatccatcgttcgcgcagatgttca	2497
Db	4870	AAGGAATATGTACATTGTTGGTAGGTTGTATTCACATATCCACCCGACCGGACATCCAC	4929
Qy	2498	tctcctcgttgcttaccctcagatattgagcgcgagaagaacagtagcatttgytggggaca	2557
Db	4930	TGCTTTACGGGACCTGAGCCTTGAGTCAAAAGGGCCAGACGCTGGCTTTGGTGGGACGA	4989
Qy	2558	gcgcgcttgggaagaagaaacttcgttcacacttcctcaagaaacttatgaaacccgtgcaag	2617
Db	4990	GTCGCTGGGGAAAGACACAGTGTGCTCAAGTCTCTGGAGCGGTTCTACAGACCCCTTTGGCAG	5049
Qy	2618	gacacgtccttcttgaatggtgtgtgacatgcacaaagaaatlgaattacagtgctccgttccc	2677
Db	5050	GGAAAGTCTCTCTTGATGAGGCAAAAGAAATTAAGGACTGTAATGTTCAHGCTCCGAGCAC	5109
Qy	2678	aaatagcaatcgttctcctcaagaagcctgtgccttccaactgtcgaacatctgcgaagaatcg	2737
Db	5110	ACCTGGGCGATGTTGCCACAGAGCCCATCTCTTTTGACATGCGACATTTCTCTAGAAATTG	5169
Qy	2738	cctatgtgtgaanaagacgcgtgtgtgtccattagatagatcgaagaagccgcgaatcgag	2797
Db	5170	CCTATGGAGACAAAGCGCGGGGTGGTGTACAGGACAGATCTGTGAGGGCCACAAAGGACGG	5229
Qy	2798	caaatatccatcttcttatttgaagtgctccctgagaaatatacacaacagaatgtgactbga	2857
Db	5230	CCAAATATCATGCTTCATCGAGTCACTGCTTAATAATATATGACCTAAAGTAGAGACA	5289
Qy	2858	aagagacacagcttctcgcgcgcagaagaacaaagactatctatctgcaaggctcttccc	2917
Db	5290	AAGGAATCAGACTCTCTGTTGGTGGCCAGAAACAAAGCATTCATACCTGCTGGCCTTTGTTA	5349
Qy	2918	aaaaaaccaaaatttatttlttgaatgagccactcagccctcgataatgacagtgaga	2977
Db	5350	GACACCTCATATTTTGTCTTTTGGATGTAAAGCCAGTCACTGTGGATACAGAAAGTGA	5409
Qy	2978	aagtgttcacagcatgccttgaataagccagagacgysgaagaacatgctagtgtgctactc	3037
Db	5410	AGCTTGTCCAGAAAGCCCTGGACAAACCCAGAGAGGCCGACCTGCATGTACAGAAAGTGA	5469
Qy	3038	acaagctctctgcgaatcattcagaagcagattgtatgtgttctgcacaatgtgaaagataa	3097
Db	5470	ACCGCGTCTCCACATCCAGATCCAGATTCACACTTAATAGTGGTTCAGAAATGGCACAGTCA	5529
Qy	3098	aggaacaaggaactcatcaagaagctcctggaataatcgaaactatatttaagttagtga	3157
Db	5530	AGGACCATATGGCAGCGATGACGCTGCTGGCACAGAAAGGACATCTATTTTTCATATGTCA	5589
Qy	3158	atgcacagtcaag	3169
Db	5590	GTTGTCACAGGCTG	5601
RESULT 13			
LOCUS	AR051647	4264 bp	DNA
DEFINITION	Sequence 1 from patent US 5630697.		PAT
ACCESSION	AR051647		29-SEP-1999
VERSION	AR051647.1	GI:5975011	

RESULT	13
LOCUS	AR051647
DEFINITION	Sequence 1 from patent US 5830697.
ACCESSION	AR051647
VERSION	AR051647.1 GI:5975011
	DNA PAT
	29-SEP-1999









Db	3773	AGGTTGTC	AAAGAGACCCCTTGGACAAAGCCACAGAAAGCCGACCTGCATTGTGATTGCTC	3832
Qy	3038	acaagctctctgcgaatc	agaacgcagattgatagtggtctctgcacaatggaagatca	3097
Db	3833	ACCGCTGTCCACCACTTCAGAAATGCAGACTTAATATGTGTGTTTCAGAAATGCAGACATCA	3892	
Qy	3098	aggaacaagaacatcatcaagaagctctcgagaacatcgacacatatatttaagtagtga	3157	
Db	3893	AGGACCAATGGCCACGCATGACAGCTGTGGCCACAGAAAGCATCTATTTTTCAAATGTCAC	3952	
Qy	3158	atgcacagtcaag	3169	
Db	3953	GTGTCCAGGCTG	3964	
RESULT	15			
LOCUS	E02326	4378 bp	RNA	PAT 29-SEP-1997
DEFINITION	Multi drug resistance relating gene derived from human normal cells.			
ACCESSION	E02326			
VERSION	E02326.1	GI:2170561		
KEYWORDS	JP 1990100680-A/1.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	1 (bases 1 to 4378)			
JOURNAL	Ueda, K. and Komano, T.			
COMMENT	HUMAN NORMAL CELL-DERIVED MDR RELATED GENE			
	Patent: JP 1990100680-A 1 12-APR-1990;			
	SUNTORY LTD			
	OS Homo sapiens			
	PN JP 1990100680-A/1			
	PD 12-APR-1990			
	PF 05-OCT-1988 JP 1988251475			
	PI UEDA KAZUMITSU, KOMANO TORU			
	PC C12N15/12, C12N1/21, C1Q1/68;			
	CC strandedness: Single;			
	CC topology: linear;			
	CC *source: tissue: Adrenal gland;			
	CC *source: clone: SAM1132;			
	EH Key	Location/Qualifiers		
	FT 5'UTR	1..137		
	FT CDS	138..3960		
	FT mat_peptide	138..3977	/gene='Multidrug resistance relating gene'	FT
	FT	/gene='MDRA1'		
	FT 3'UTR	3981..>4378.		
	FEATURES	Location/Qualifiers		
	source	1..4378		
		/organism='Homo sapiens'		
		/db_xref='taxon:9606'		
	BASE COUNT	1315 a 818 c 1057 g 1188 t		
	ORIGIN			
	Query Match	35.3%; Score 1122.4; DB: 6; Length 4378;		
	Best Local Similarity	62.0%; Pred. No. 9, 9e-221;		
	Matches 1868; Conservative	0; Mismatches 1096; Indels 48; Gaps 4		
Qy	200	tttaagatgatacaagaatc	atcaagaatgacaagaatttgacataaagaagacatag	259
Db	958	TTTAAAGGATCAACAAATTTTGAAGAAGCTTAAAGAAATTGGGATTAAGAAAGCTATTA	1017	
Qy	260	cttcaaaaatgctctctgtgctgtgttactctttagaataaactatgagactgtctt	319	
Db	1018	CACCCAAATATTTTATAGAGGCGCTTCTCTGCTGATTCATCATCTTAATGCTGGCCCT	1077	
Qy	320	tttgataggaacccctcccttgatctttaaityggaacccatgataaccatcggaactgttc	379	

Db	1078	TCCTGTATGGGACACCTTGCTCTTCAGGGGA-----ATATTCTATTGGACAAGTAC	1131
Qy	380	ttgcgtcttcctcttagtgcgtatccatagcagttatgcattcgtatcgtgagcaagctccctact	439
Db	1132	TCACCTGATTCTCTTTTCGTATTAAATTGGGGCTTTAGTGTGGACAGGCATCTCCAGCA	1191
Qy	440	ttgaaccttcgcataagcccgagagctgtcccttcataatttccaggttatctgataaga	499
Db	1192	TTGGAAGCATTTGGCAAATGCAMAGAGGAGCAGCTTATGMAATCTTCAAGTAATTTGATATA	1251
Qy	500	aaccagatagataaactttccacagctcgtgatatanaacctcgtaacctcagtagaagactg	559
Db	1252	AGCCAAAGTATTGACAGCTATTTCGAAGAGTGGCACAAACCAATATATTTAAGGGAAATT	1311
Qy	560	tggaatttaaaaaagttcttcctcaatcatcocalcaagaccatctcaagaattctgaag	619
Db	1312	TGGAAITTCAGAAATGTTCACTTCAGTTACCCATCTCGAAMAGATTTAAGATCTTGAAG	1371
Qy	620	gtctgaatctcagaattaagltctgagagagacagtcgtcccttgcgttcgtctcaatgcag	679
Db	1372	GTCGCAACCTTGAAGGTCCAGAGTGGCAGCGGTGGCCCTGGTTGGAAACAGTGGCTGTG	1431
Qy	680	ggaagaagcagtgatgcagctcttcacagaggtatagatcgcgatgatgattcatca	739
Db	1432	GGAAAGCACACACAGCTCCAGCTAGTGCAGAGGCTCTTATACCCACACAGGGAGATGCTCA	1491
Qy	740	tggtgcagagaaatgcatacagaacttlaaagtctgcgcatatcgcagacataltgag	799
Db	1492	GTGTTGATGCACAGATATTATAGGCCATAAATGTAAGTTTCTACGGGAAATCATCTGGTG	1551
Qy	800	tggttagtcaagagcctgttttcttgaggagcccaatcagtaacatatcaagtatgag	859
Db	1552	TGGTGAGTCAGAAACCTGTATTGTTGGCCACCCAGTACCTAATAAACATCGCTATGAGCC	1611
Qy	860	gagatgcagtgcagatgcagaagagatgagagagcagaaggaacaaatgcgtatgatt	919
Db	1612	GTCAAATGATTACCATGAGGATGAGATTGAGAAAGCTGTCAAGAGAGCAATGCTTACT	1671
Qy	920	ttatcatcagagtttccctaataatttaatacatctgtgtaaggggaaaaagagctcaatga	979
Db	1672	TTTATCATGAACCTGCTCATTAATTTGACACCTGGTTGAGAGAGAGGGGCCAGTTGA	1731
Qy	980	gtgcagggcagaaacagagatgcgaatctgtctgtccttagttcgaaaccccaagatc	1039
Db	1732	GTGGTGGGCAACAGCAGAGGTCCCTATTCACAGTCCCTGGTGTGCTAACCCCAAGATCC	1791
Qy	1040	tgattttgagtgagagctcagctcctcccttgattccagaagaagatcagctgtctcaagctg	1099
Db	1792	TCTCTCTGTGATGAGGCCAGCTGACCTCTTGAGCACAGAAACGAAAGCAGATGGTTCAGGTGG	1851
Qy	1100	caactgcagaaagcgcagcaaaagltgcgactcaactcgtgtgtagcacacgcgacttcta	1159
Db	1852	CTCTGTGATTAAGGCCAAGAAAGTGGGACCACTTGTGATGATGCTCATCTTGTGTACAG	1911
Qy	1160	ttcgaagtgcagatttgattgtgcacctaaagatgtgaatgctcgtgcgaggaagaagcac	1219
Db	1912	TTTCGTAAATGCTGACGTCTATTCCTGTTTCATATGAGATGAGATTTGGGAGAAAGAAATC	1971
Qy	1220	atgcgtcaactaattgcaaaagaggtctatattatcaactcgtgattgcgtacaggaatata	1279
Db	1972	ATGATGAATCATCTAGCAAGAAAGAAAGCATTTACTTTCAAACTGTGTCAAAATGCAGACAGCG	2031
Qy	1280	aaaaagctgcatagaagaatgtagtcaatlgacatatcta-----ctga	1322
Db	2032	GAAATGAAAGTTGAATTAGAAAATGCACGTGATGATTCACAAAGTGAATTGATGCTTTGG	2091
Qy	1323	aagaaagaccaactcactctcctgcacctctgtgaaagacata-----	1366
Db	2092	AAATGTTCTTCAATGATTCAGAGATCCAGCTGTAATATAGAAAAAAGATCAACTCGTAGAGTG	2151
Qy	1367	-----agtcagactcatctbaacaagctgcaggaatccaccacaaactaaagaatc	1420
Db	2152	TCCGTGGATCAACACCCCAACACAGAAAGCTTACTATCCAAAGAGGCTCTGTGATGAAGTA	2211

Oy 1421 ttccctgaagctctctctatataaaatttaagaattaaacagctgaatgacctttgtg 1480  
 Db 2212 TACCCGAGGTTTCCTTTGGAGGATTAAGCTAAATTAACCTGAAGGGCTTTATTTTG 2271  
 Oy 1481 ttctggagacatttgctctcgtcctaaatggaaactgtccatccagtatcttccacatct 1540  
 Db 2272 TTGTTGGTATTTTGTGCTCANTATATAATGGAGGCTGCAMCAGCATTTTGCATATAT 2331  
 Oy 1541 ttgcataaattataacacatglttggaaaataatgataaaacacatlaag---catgag 1597  
 Db 2332 TTTCAAAATATATAGGGGTTTTTACAAAGATTGATCTGAAACAAACAGCAGATA 2391  
 Oy 1598 cagaaattatccatglatcgcatttgggtgtgtatttggcttggctgttccatcc 1657  
 Db 2392 GTAACTTTTTCCTATTTGTTCTAGCCCTTGGAAATTAATTTCTTTATTTATTTTC 2451  
 Oy 1658 tgcaggaattatttaagcagagcagaggaattttaagatgaatgaactgtg 1717  
 Db 2452 TTACAGGTTTCANTTTGGCAAGCTGAGAGATCTTCACCAAGCGGCTCCGATTCATGG 2511  
 Oy 1718 ccttcaaaagcattatatacagaatltgcctggttggatgaagaaagaaacagacag 1777  
 Db 2512 TTTTCCGATCCATGCTCAGACAGATGTGAGTGGTTGATGACCCCTAAACACACATCG 2571  
 Oy 1778 gaagcttgaacaataatataagcacaataatcaagagcaagacagttcca 1837  
 Db 2572 GAGCTTTACATACCAAGGCTCCCAATGCTGCTCAAGTTAAAGGGCTTAATGTTTCCA 2631  
 Oy 1838 gaattgagcttcaacaataatgaactaaacatgaaggaacttccagttaacttcc 1897  
 Db 2632 GGCTTGGCTATATTTACCAAGATATAGCAAACTTGGGACAGGAATATATATCTTCA 2691  
 Oy 1898 tatalgagttggagatgaactcctgattctgagatagttcagtaactgcgtgacag 1957  
 Db 2692 TCTATGTTGGCACTACACTGTTACTCTTATGCAATTTGATCCCATTTGCAATAGCAG 2751  
 Oy 1958 gaattgagaaacagcagcaatgactgatttgcacaacaagaatgaagaacttaagc 2017  
 Db 2752 GAGTTGTTGAATGAATAATGTTGTGCAACAGCACTGAATGAATGAATAAGATAGAG 2811  
 Oy 2018 atgcttgaagaatagaactgaagcttggagaatatacgtactaagtgacttaacaa 2077  
 Db 2812 GTTCTGGAGATGCTCTGATGAAGCAATAGAAACTTCGAAACGTTGTTCTTGACAC 2871  
 Oy 2078 gggaaaagccttcagcaaatgtaagaagatgcttcagacccaacaaataact 2137  
 Db 2872 AGGACAGAAAGTTGAAATATGTATGCTCAGAGTTTGACAGGTACATACAGAACTCT 2931  
 Oy 2138 cgaagaagaacagatlttgaagcgttactgactcaagcctgacttataatlttg 2197  
 Db 2932 TGAGGAAAGCACACATCTTGGAAATTAACATTTCTCCACCCAGCAATGATGATTTT 2991  
 Oy 2198 cctatgcagcagaggttgcatttggagcctatlttaacgaagctgaagaaatgacccag 2257  
 Db 2992 CCTATGCTGATGTTCCGGTTTGAGAGCTACTTGTGGCAATATAACTCATGAGCTTTG 3051  
 Oy 2258 agggcagtgatcatgatttactcgtcaatgcaatgagcctgagcaatcggaanaacgc 2317  
 Db 3052 AGGATGTTCTGTATGATTTTACAGCTGTGTGCTTGTGCTATGGCGTGGGCAAGTCA 3111  
 Oy 2318 tgcgttggctccctgaatatatccaaagcaaacgcggggcgtgcgcatcgttgcctgt 2377  
 Db 3112 GTTCAATTTGCTCTGACTAGTCCAAAGCCAAATATCAGACGCCACATCATCATGATCA 3171  
 Oy 2378 tggaaaagaaccaaatalagacagccgagtcgaagaaaggaagaaacagacacatgtg 2437  
 Db 3172 TTGAAAAAACCCCTTGAATGACACACTACAGCAGGAAGGCTTAATGCCGAACACATTTG 3231  
 Oy 2438 aagggaatttagagtttcgagaagtcctcttctatccatgctgcgcagatgtttcca 2497  
 Db 3232 AAGGAATGTCAATTTGGTGAAGTTGATTCACATATCCACCGACCGGACATCCAG 3291

Oy 2498 ttccctgagcttaccctccaglatgagcaggaagacagatgacatttggggagca 2557  
 Db 3292 TccttCAGGAGCTAGGCTCGAGGTGAAGAGGCCAGACGCTGCTGTGGGGCAGCA 3351  
 Oy 2558 ggcgttgggaaaagacacttctgtcacttctgaagacatttgaacccgtgacag 2617  
 Db 3352 GTGGCTGTGGGAAGACAGACAGTGTCCAGCTCCTGGAGGGTGTACGACCCCTTGGCAG 3411  
 Oy 2618 gacaagtgctgttgaatggttgaatgcaaaaagaattgaatgacatggtgcgtcc 2677  
 Db 3412 GGAAGTCTGCTGTGATGGCAAAAGAAATGAACCGACTAATGTTTCAGTGGCTCCAGAC 3471  
 Oy 2678 aatagcaatcttcccttaagaagcctgtgccttcaactgaagcaatgtgaagaacatcg 2737  
 Db 3472 ACCTGGCATGCTGTCCAGAGGCCATCTGTTGATGACGACATTTGCTGAACATTTG 3531  
 Oy 2738 cctatgtaacaacagccgttgtgtccattgaatgaatgaagaagccgcaaatgag 2797  
 Db 3532 CCTATGAGAGACAGCCGGGTGTGTCAAGGAAGATTTGTGAGGCGACGAAGAGAG 3591  
 Oy 2798 caaatatccattcttattgaaggtctccctgagaatatacaacaacagatltgactga 2857  
 Db 3592 CCAACATACATGCCCTTCATCGAGTCACTGCTTAATTAATATAGCACTAAGATGAGACA 3651  
 Oy 2858 aagagcagcgttcttctgtgcggccagaaagaagactgacttgaagggctctcc 2917  
 Db 3652 AAGGAATCAGCTCTGTGTGTGGCCAGAAACAGCATTTGCACTGCTGCTGCTTGTGA 3711  
 Oy 2918 aaaaaccaaataatttattgttgaatgaagccacttgcgacctgataatgaatga 2977  
 Db 3712 GACACCTCATATTTTGTGCTTTTGGATGAAGCCAGCTGAGCTGTGATACAGAAATGAAA 3771  
 Oy 2978 aggtgttcagcatgccccttgaataagcaggaaggaagacatgactagtgactc 3037  
 Db 3772 AGGTGTCCAGAAAGCCCTGGACAAAGCCAGAGAGGCCGACCTGCATTTGATTTGCTC 3831  
 Oy 3038 acagcctctcgaattcagaagcagatltgatatggttgcgttcacaaatgaagataa 3097  
 Db 3832 ACCGCTGTCCACCTCATGAAATGCAAGCTTATATGTGTGTTTGAATGCAAGTCA 3891  
 Oy 3098 aggaacaaagaaactcaagaagcctccttgaagaatcgaacatalatlttaagtgtga 3157  
 Db 3892 AGGACATGGGACAGCATCAGCAGCTGTGCGCACAGAAAGCATATTTTCAATGTGCA 3951  
 Oy 3158 atgcagatcag 3169  
 Db 3952 GTGTCCAGGCTG 3963

Search completed: April 22, 2002, 22:48:25  
 Job time: 27413 sec



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:41:29 ; Search time 293.81 Seconds

(Without alignments)  
2853.618 Million cell updates/sec

Title: US-09-873-409-13

Perfect score: 3702  
1 ttcgcgttcgcgtgactgact.....atgcacagtcacgtgcgtga 3702

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/6C.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/Backfile1.seq: \*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1225.2	33.1	4646	1 US-08-181-471-2	Sequence 2, Appli
2	1225.2	33.1	6505	2 US-08-793-610-5	Sequence 5, Appli
3	1225.2	33.1	9318	2 US-08-793-610-6	Sequence 6, Appli
4	1223.6	32.9	4669	6 5206352-3	Patent No. 5206352
5	1217.2	32.9	4669	2 US-08-583-276-18	Sequence 18, Appli
6	1217.2	32.9	4669	2 US-08-752-447-1	Sequence 1, Appli
7	1214.2	32.8	4264	2 US-08-784-649A-1	Sequence 5, Appli
8	1212.6	32.8	4264	2 US-08-784-649A-5	Sequence 1, Appli
9	1121	30.3	4233	3 US-09-120-513-1	Sequence 1, Appli
10	1121	30.3	4233	4 US-09-450-105-1	Sequence 1, Appli
11	856	23.1	2726	1 US-08-461-823-1	Sequence 1, Appli
12	426.6	11.5	4002	2 US-08-996-545-3	Sequence 1, Appli
13	426.6	11.5	4002	2 US-08-996-545-3	Sequence 3, Appli
14	426.6	11.5	4002	4 US-09-328-320-1	Sequence 1, Appli
15	426.6	11.5	4002	4 US-09-328-320-3	Sequence 3, Appli
16	386.6	10.4	4047	2 US-08-612-734B-1	Sequence 1, Appli
17	331.4	9.0	4224	1 US-08-612-521-1	Sequence 1, Appli
18	314.6	8.5	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	304.6	8.2	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	193.2	5.2	3927	2 US-08-996-644-1	Sequence 1, Appli
21	193.2	5.2	3927	3 US-09-352-552-1	Sequence 1, Appli
22	192.4	5.2	3924	2 US-08-996-644-3	Sequence 3, Appli
23	192.4	5.2	3924	2 US-09-352-552-3	Sequence 3, Appli
24	181.6	4.9	3909	1 US-08-232-537-1	Sequence 17, Appli
25	149	3.9	2061	4 US-09-061-764A-17	Sequence 2, Appli
26	145.2	3.9	2407	2 US-08-895-522-2	Sequence 2, Appli
27	145.2	3.9	2407	3 US-09-195-391-2	Sequence 2, Appli

28	145	3.9	5120	3 US-08-772-270A-6	Sequence 6, Appli
29	145	3.9	8370	2 US-08-488-706-1	Sequence 1, Appli
30	144.8	3.9	7721	3 US-08-772-270A-14	Sequence 14, Appli
31	141.2	3.8	1959	4 US-09-061-764A-4	Sequence 4, Appli
32	133.4	3.6	1201	3 US-08-961-083-189	Sequence 189, App
33	133	3.6	2244	4 US-09-061-764A-18	Sequence 2, Appli
34	130.4	3.5	4403765	4 US-09-103-840A-2	Sequence 1, Appli
35	128.6	3.5	2376	1 US-08-394-880B-1	Sequence 2, Appli
36	128	3.5	4403765	4 US-09-103-840A-2	Sequence 1, Appli
37	128	3.5	4411529	4 US-09-103-840A-1	Sequence 1, Appli
38	117.6	3.2	6143	1 US-08-612-521-3	Sequence 3, Appli
39	110	3.0	28804	2 US-08-592-874-1	Sequence 2, Appli
40	110	3.0	28804	3 US-09-096-942-2	Sequence 2, Appli
41	110	3.0	28804	3 US-09-096-942-2	Sequence 2, Appli
42	104.2	2.8	5889	1 US-08-463-092B-5	Sequence 5, Appli
43	104.2	2.8	5889	2 US-08-462-109A-5	Sequence 5, Appli
44	104.2	2.8	5889	2 US-08-460-907B-5	Sequence 5, Appli
45	104.2	2.8	5889	3 US-08-463-179A-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-08-181-471-2  
Sequence 2, Application US/08181471  
Patent No. 5641508  
GENERAL INFORMATION:  
APPLICANT: Li, Lingna  
INVENTOR: Lishko, Valery K.  
TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Thomas Fitting  
STREET: 12526 High Bluff Drive, Suite 300  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92130  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 13-JAN-1994  
CLASSIFICATION: 424  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: ANT0029P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-792-3680  
TELEFAX: 619-792-8477  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4267  
US-08-181-471-2

Query Match	33.1%;	Score 1225.2;	DB 1;	Length 4646;
Best Local Similarity	60.9%;	Pred. No. 0;		
Matches 2178;	Conservative	0;	Mismatches 1268;	Indels 132;
				Gaps 6;

QY	162	caggcttgcacccctgattatgcttgtaataagggtgctgctccttgatttcttgctgatactaa	221
Db	760	CAGGATGAGCCATATTATACAGTGGAGATTGGTGGCTGGGCTGGTGGCTGCTTACATTTCA	819
QY	222	gatttcctctgtggtatcaactcgcgcagcagcagcagcagagagattccgaacagtttct	281
Db	820	GCTTTTCATTTTGGTGGCCCTGGCGACGTGGAAAGCAAAATTAACAATAATTAGAAAACTTTT	879
QY	282	tcaatcagtttttgcgcagagacacgcgctgtgttatagtcgtlgaacatcgltgaactaa	341
Db	880	TCATGCTATAATGGCACAGGAGATGAGCTGGTTGATGTGCACGATGTTGGGGACCTTAA	939
QY	342	cactcgcgttcgcagacatctgac---aaatacgaatgagtatttgataagaattgacct	398
Db	940	CACCCGACTTTCACAGTGAATGTCTTAATAATTAATTAAGTTATGTGTGCACAAATTGGAA	999
QY	399	gtgtttcgaacaaacatgctacttcttcgatctggccctggcagltggttttggltgaagagctg	458
Db	1000	GTTCTTTCATGCAATGGCAACATTTTTCACGTGGGTTTATAGTAGGATTTTACACGTGGTTG	1059
QY	459	gaaactcaccttagtgaactatcacagctctcctcttaatagtcttaagcgagcagatg	518
Db	1060	GAAGCTAACCCCTTGATTTGGCCATGACCTGTTTGGACATGTCAGCTGCTGTG	1119
QY	519	ttctcgaagtgctcatctatcgacagcaagaatlaagtgctatcttcaaacgtggagc	578
Db	1120	GGCAAGATTAATCTTCTTATTTACGATTAAGAACTCTTACGGTATGCAAAAAGCTGGAC	1179
QY	579	tgtagcagaagaagctctgtgcatcaalcgcagacagtcacgtcttaagggcccgagaa	638
Db	1180	AGTAGCTGAAGAAGGCTTGGCAGCACTTATTAAGAACTGTGATGGATTTGGAGCAAAAG	1239
QY	639	agaactccaagggtcttcctttaataatacaagaatagctgtgtttatttcccca	698
Db	1240	AGAAC-----	1244
QY	699	gtggtactaagtgctgtcttglncttgtaaggtatacacagaatctcaaaagtgcata	758
Db	1245	-----TTGAAGGTACAAACAAATTTAAGAAAGACTTAA	1278
QY	759	ggattttgcatataaagaagcatatagcttcaaaagtgctctgtgctgtgactctct	818
Db	1279	AAGATTGGGATTAAGAAAGCTATTATACGCCAATTTTCTATAGGGTGGCTTCTCTGCT	1338
QY	819	tatgaatgaacctaagtcactgtccttttggatgaagaacctcgtattcttaatgagaa	878
Db	1339	GATCATATCATCTTATGCTCTGGCCTTCTGGTATGGGACACACTTGGTCTCTCAAGGGGA	1398
QY	879	acctggaataccatcggagactgtctgtcgtgttctctttagtgaatcaatagacagta	938
Db	1399	-----ATATCTATTATGGACAAGTACTCACTGATTTCTTTTCTGTATTAATGGGGCTTT	1452
QY	939	ttgcatctggagcagcagctccctcactcttggaaacctgcgaatagcccgagagagctgctt	998
Db	1453	TAGTGTGGAGCAGGCACTCTCCAGAGCATTTGAAGCAATTTGCAATGCAAGGGGACAGCTTA	1512
QY	999	tcaattttcagagctatctgataagaacccagatagataactcttccacagctcgata	1058
Db	1513	TGMAATCTTCAAGATTAATTGATTAATTAAGCCCAAGTATGACACTTATTTGAAAGATGGGCA	1572
QY	1059	taaacctgaatccatagaagaagactgttggaaattlaaaaatgcttcttccaattacatc	1118
Db	1573	CAAAACAGATTAATTAAGGGAATTTGGAATTTGGAATTTGCAATGTTCACTTCAGTTACCAATC	1632
QY	1119	aagacatcatatcaaatcttgcgaagagctgcgaatctcgaatlaagctcgtgagagacgt	1178
Db	1633	TCGAAAACAAAGTTAAGATCTTGAAGAGGCTGAACTGGAAGGTGCAGAGTGGGCAACAGCT	1692

QY	1179	cgcttgctgctgcctcaatgycagtgaggaaagcagctagtcacgtctctgcggaggtc	1238
Db	1693	GGCCCTGGTTGGAAACAATGGCTGTGGAAAGACAAACAATCCAGCTGATGAGAGGCT	1732
QY	1239	atcgtatccggatcgtatgctcttcaatcgtgtgacgtatgatatcaacagctcttaaatgt	1298
Db	1753	CTATGACCCCAACAGAGGGGATGCTACAGTTGATGAGACAGGATATTGACACATTAATGT	1812
QY	1299	gcggacatcgcagacaacatltgagtgagtgatgtacgaagcctgttctgtcggagccac	1358
Db	1813	AAGGTTTCTACGGGAAATCATTTGGTGTGTGAGTCAGAGAACCTGTATTGTTGGCCACAC	1872
QY	1359	catcagtaacatatloaagtaatlgacagagatgcatgctgcatgaaagatgagagagac	1418
Db	1873	GATAGCTGAAAAACATTCGCTATGGCCGTGAAATAATGTCCACATGATGAGATTGGAAAGC	1932
QY	1419	agcaagaggaagcaaatgctgatagtatcttcatcgtgagttccctaataatcaatc	1478
Db	1933	TGTCAGGAAACCCAAATGCTTATGACTTTATATGAAATGCTCTATTAATTTGACACCCT	1992
QY	1479	ggtlagggaaaaagagcctcaaatlgatgltggaagggcagaaacagagatcgcgaattgtcg	1538
Db	1993	GGTTGGAGAGAGAGGGGCCAGTTGAGTGTGAGTGGGAGAGACAGAGATGCCATTGACAG	2052
QY	1539	tgccctagttcgtgaaaccccaagatctgtaatttgatgtaggtctacgctccctggatlc	1598
Db	2053	TGCCCTGGTTGGCAACCCCAAGATCTCTCTCGATGATGAGCCAGTCAGGCTTTGGACAC	2112
QY	1599	agaaagcaagtcacagctgtctcaagctgcacctgaggaaggcagacaaagctgcgactacaat	1658
Db	2113	AGAAAGCCAAACAGTGGTTCAAGTGGCTCTGTGATNAGGCCAGAAAAGGTCCGACCACCAT	2172
QY	1659	cgtlgtlagaacacagcactcttctactatcgaagtcgaagatttgaattgtgacccctaagga	1718
Db	2173	TGTGATACCTCATGCTGTTGTCTACAGTTCGTAATGCTGACGTACGCTGGTTCGATGA	2232
QY	1719	tggaatgctgcgcggaagaaagagacacatgctgcgaactaatgycaaacagaggtctcatatla	1778
Db	2233	TGGAGTCATTTGTGGAGAAAGAAATCATGATGAATCTCATGAAAGAAAGGCACTTTACTT	2292
QY	1779	ctcactgtgtgctgcacagatatlaaaaagccgatatcaagatgagatcgaatcagacata	1838
Db	2293	CAAACTGTCTCAATGCAACAGACAGGAAATGAAGTTGAATTTAGAAAATGCACGTGATGA	2352
QY	1839	ttctca-----ctgaaagaaagaccaaactcactccctcgtcactcgtg	1881
Db	2353	ATCCAAAAGTGAAATGTGATGCTTGGAAATGTCCTTCAAAATGATTCAGATCCAGCTCAT	2412
QY	1882	aagagcatca-----agtcagactctcattgaacagctcgaga	1919
Db	2413	AAGAAAAAGATCAACTCTAGAGGTGTCCGTGGATTCACAMCCCAACAAAGAAAGCTTAC	2472
QY	1920	atcacccaatctaaagagataaagatcttcctccgaagtcctcetatlaaaatctttaaagtt	1979
Db	2473	TACCAAGAGGCTCTGGGTGAAGATNACCTTCAGATTTCTTTTGGAGGATTATGAACCT	2532
QY	1980	aaacaagcctbaatgagccttltgtgttctcgtgaggaacatttggtcctctgtctcaaatggaac	2039
Db	2533	AAATTTAACTGAATGGCCTTATTTTGTGTGTGGTGTGATTTTGTGCCATTAATAATGAGAG	2592
QY	2040	tgttcatcaglatlctccalcactcttgcaaaataataaacabgltt--ggaataaa	2096
Db	2593	CCTGCACACACATTTGCAATAAATTTTCAAGATTTATAGGGGTTTATTACMAAATTTGA	2652
QY	2097	tgataaaacccacatlaagcagatgacgaagaattatctccatgatatcgtcatttggg	2156
Db	2653	TGAATCCGAAACAAACACAGCAAGATAGTAACCTTGTTCCTACTATTGTTTCAAGCCCTTGG	2712
QY	2157	tgattatggtctgtcagtiaattccaatgcaggaattattcaagcgaagcaggggaaat	2216
Db	2713	AATATATTTCTTTTATACATTTTTCCTTCAGAGGTTTTCACATTTTGGCAAAAGCTGGAGAGAT	2772
QY	2217	tttaacagctgatatgaacacttgccttcaaacgcatgltatatacagatatcgcgtg	2276

Db 2773 CTTCCAGACGGCTCGATATCATGTTTCCGATCCATGCTCAGACAGATGTGAGTTG 2832  
OY 2277 gtttgaatgaagaaagaaacagacagagagcttgaacaacatattagccatagatagc 2336  
Db 2833 GTTGTGATGACCCCTAAACACACACTGAGACTGATACCAAGCTCCGCAATGATGCTGC 2892  
OY 2237 acaaatccaaggagcaacaggttccagatttgcgtcttcaacaacaaatgaactaacat 2336  
Db 2893 TCAAGTTAAAGGGGCTATGATGCTCCAGGCTTGCTGTAATTAACCAAGATATATGAAACT 2952  
OY 2297 ggaacttcagttacatttccttataatataatgagagatgacatcttcagttcag 2456  
Db 2293 TGGACACGAGATATATATATCTTCATCTATGTTGGCAATACACTGTTACTTAC 3012  
OY 2457 taattgcagtagtctgcgtgacaggaatgattgaacccagcaaatgactgtattgc 2516  
Db 3013 AATTGTATCCCATCATCTCAATAGCAGAGTTGTAATGAAATGTAATGTTGCTGGAACAAGC 3072  
OY 2517 caacaaagataagcaagaacttaagcaltgctggaagaatagcaactgaagcttggagaa 2576  
Db 3073 ACTGAAAGATTAAGAAAGAACTAGAGAGGCTGGGAGAGATCCCTACTGAAAGCAATAGAAA 3132  
OY 2577 tatacgtactatagtgcatataaacaagaaagcccttcgagcaaatgtaagagat 2636  
Db 3133 CTTCCGAAACCGTTGTTCTTTGACTCAGAGACACAGATTGTAACATATGTAATGCTCAGAG 3192  
OY 2637 gcttcaagactcaacaagaactaacctcgaagaagaacagatlaattggaagctgtatgc 2696  
Db 3193 TTTGACAGATACCATACAGAAACCTTTGAGAGAAACACACTTTTGGAAATTAATTTT 3252  
OY 2697 attcaagcagcttataatatttgcctatgcagaggttccagatttgcagctattt 2736  
Db 3253 CTTACACAGCAATGATGATGATTTTCTATGCTGAGATGTTCCGGTGTGGAGGCTACTT 3312  
OY 2757 aatcaagcttgaagaaatgagcccaagagagcaltgcatatlaatttactgaattgata 2816  
Db 3313 GGTGGACATTAATCAATGAGCTTGGAGATGTTCTGTAGATATTTTCAAGCTGTGTT 3372  
OY 2817 tgaagcaltgagcaltggaagaaacgctcgttcttgcctcgtgaatattccaaagcaatc 2876  
Db 3373 TGGTGCATGGCCGTGGGGCAAGTCAAGTTCTGTTGCTGACTGACCAAGCAAAAT 3432  
OY 2877 gggggcgtgcgcatctgttgcctgttggaaagaacaaatataagacagccgagctca 2936  
Db 3433 ATCAGCAGCCCATCATCATGATCATGATAAACCCTTGTGATGACACACTACAGCAC 3492  
OY 2937 agaagggaaaaaagcagacagacatgtgaaggaattagagttcgaagaagctcttctt 2996  
Db 3493 GGAAGGCTTATGCGCAACATTTGGAAGAAATGTCAATTTGTTGAAGTTGATTTCA 3552  
OY 2997 ctatccatgtgcagcagatgtttcattcctcgttgcgttaccctcagttatggagag 3056  
Db 3553 CTATCCACCCGAGCCGAGACATCCAGTGTCTCAGGACTGAGCTGGAGAGGAAAGAGG 3612  
OY 3057 aaagacagtgcatcttgttggagagcagcgtgtggaagaaagccttctgtcaacttct 3116  
Db 3613 CCAACACCTGGCTGTGGGACACACTGTGCTGGAAGAGACAGTGTCCAGCTCCT 3672  
OY 3117 gcaagacattatgaccccggtgcaaggaagtgctgttattgtgtgagtgcaaga 3176  
Db 3673 GGAGCGGTTTACGACCCCTTTGGAGGAAAGTCTGCTTATGCAAGAAATTAAGCG 3732  
OY 3177 attgaatgtacagtgtgctcgttcccaaatagcaatcgttccctcaagagcctgtgctt 3236  
Db 3733 ACTGAATGTTCAAGGCTCCGAGCACACTGGCATCTGTCAGAGAGCCCATCTCTT 3792  
OY 3237 caactgaagcattgtgtgaacacatcgctatagtgagacacagcgtgtgtggtcattaga 3296  
Db 3793 TGACTGACGACTTGTCTGAAGCAATTTGCTATGAGACAAACGCGGAGTGTCAACAGA 3852  
OY 3297 tgaagatcaagaagccgcaaatgagcaaatatcatcttatttgaaggttcccttga 3356

Db 3853 AGAGATGCTGAGGGCAGCAAGAGAGGCCACATACATGCTTCAATGAGATCTGCTTAA 3912  
OY 3357 gaataacacacacagttgactgaaagagcagacatttctggtgcggaagaaag 3416  
Db 3913 TAAATATAGCACTAAAGTAGAGAGACAAAGGAACCTCAGCTCTGTGCTGCGCAAAACAGC 3972  
OY 3417 actgactatgcaagggcttcttccaaaacccaatattatgttggatggagggcagc 3476  
Db 3973 CATGGCCATACCTGCTGCTGCTGTTAGACAGCCTCATATTTTGTGTTGATGAGAGCAC 4032  
OY 3477 ttccagccctcgaataatgacagtgagaaagtggttcaagatgccttgaataagcagagac 3536  
Db 4033 GTCAGCTCTGATACAGAAATGAAAGTGTTCACAAAGACCCCTGGACAAAGAGAGAG 4092  
OY 3537 ggaagagacatgcttagtgctcaactcacagctctctgcaattcaagcagagattgat 3596  
Db 4093 AGGCGGACCTGCATTTGATTTGCTCACCGCTGTGCCACATCCAGATGCAAGCTTAT 4152  
OY 3597 agtgtctgcacaaalgaaagataaaggaacaaaggaactcaatcaagagcctcagagaa 3656  
Db 4153 AGTGTGTTTCAAGATGAGCAGATCAAGAGCATGCGACGATCAGCAGCTGTGCGACA 4212  
OY 3657 tcgagacatattttaaagttagtgaaatgcacagtcag 3694  
Db 4213 GAAGGCACTATTTTCAATGATGTCAGTGTCCAGGCTG 4250

RESULT 2  
US-08-793-610-5  
: Sequence 5, Application US/08793610  
: Patent No. 5858744  
: GENERAL INFORMATION:  
: APPLICANT: BAUM, Christopher  
: APPLICANT: STOCKING-HARBERS, Carol  
: APPLICANT: OSTERLAG, Wolfram  
: TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
: TITLE OF INVENTION: FOR GENE TRANSFER  
: NUMBER OF SEQUENCES: 6  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
: STREET: 655 Fifteenth Street N.W. Suite 330  
: CITY: Washington  
: STATE: D.C.  
: COUNTRY: U.S.A.  
: ZIP: 20005-5701  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentln Release #1.0, Version #1.30.  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/793,610  
: FILING DATE: 07-MAR-1997  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE P 44 31 973.8  
: FILING DATE: 08-SEP-1994  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 195 03 952.1  
: FILING DATE: 07-FEB-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: PCT/EP95/03175  
: FILING DATE: 10-AUG-1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Beiman, Richard J.  
: REGISTRATION NUMBER: 39,105  
: REFERENCE/DOCKET NUMBER: P1614-7007  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (202)638-5000  
: TELEFAX: (202)638-4810  
: INFORMATION FOR SEO ID NO: 5:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 6505 base pairs  
: TYPE: nucleic acid







```

Oy 2097 tgataaaacacattaaagcatgtagcagaattattatccatgatattcgtcatttggg 2156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4004 TGATCTCGAACAACAAACAGCAAGAAATGTAACCTGTTTTACATATGTTGTAAGCCCTTGG 4063
Oy 2157 tctatttcttctgtagtatttctatcagcaggatatttctacagcagaggggaaat 2216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4064 AATATATTTCTTTTATATATTTTCTTCCAGGTTTCACATTTTGGCAAAAGCTGGAGAGAT 4123
Oy 2217 tttaagatgagatlaagacacttggccttcaaaagcatgtatatacagagatattgctg 2276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4124 CCTCCCAAGGCGCTCCGATACATGTTTTCGATCCATGTCACAGACAGAGATGAGTTGG 4183
Oy 2277 gtttgaatgaagaaacagacagagaggttgaacaacatatatagccatagatatgc 2336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4184 GTTGTATACCTCTTAACACACCACTGAGCANTGACTACAGGCTCGCAATAGTGTGC 4243
Oy 2337 acaaatlaagagagcaacaggttccagagattggcgtcttaacacacaaatgcaactacat 2396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4244 TCAATTTAAAGGGCTATAGGTTCCAGGCTTGCTGTATTTACCCGAATATAGCAATCT 4303
Oy 2397 gggacttcaatcattcccttataatagatggagatgagacatccctgattctgag 2456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4304 TGGGACAGGAATTAATATATCTTCTATGCTTGCGCACTAGTGTACTCTTAC 4363
Oy 2457 tatgtctcaatcttgcgcgtgacagagatattgaaccgcagcaatgactgtgttgc 2516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4364 AATTTTACCCTCATTCATTAATGACGAGGAGTGTGGAATGAAATGTTGTGGCAAGC 4423
Oy 2517 caacaaagatagcaagaacttaagcagctggaagaatagcaactgaagcttggagaa 2576
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4424 ACTGAAATATAGAAAGAAAGAAAGAGTGCTGCGAAGATCCTACTGAAAGCAATAGAAA 4483
Oy 2577 tatacgtactaagtgatcatatacaagagaaagccttcgagcaaatglatagagat 2636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4484 CTTCCGAACCGTTGTTCTTGTGACACAGAGAGCAAGTTTGAACATATATGATGTCAGAG 4543
Oy 2637 gcttagactcaacacaaataactctgaaagaacacagatatattgaaagtgtatgc 2696
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4544 TTTGAGGTACCATATACAACTCTTTAGAGAAACACACACTTTTGGAAATTAATTTTC 4603
Oy 2697 attcagcatgccttataatatttgcctatgcaagcaggttgcatttggagctattt 2756
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4604 CTTACCCAGGCAATGATGATTTTCTATCTGATGATGTTCCGTTTGGAGCTACTT 4663
Oy 2757 aattcaagctgagcaatgaccccaagagcagtgatcatagtttcttactgcatgcata 2816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4664 GGTGACATATAACTCATGAGCTTGAAGATGTCTGTATGATTTTCAAGCTGTGTCT 4723
Oy 2817 tggagctatgcccacggaaaaagcgtcttctgctcctgaatatccaaagccaatc 2876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4724 TGGTGCCATGGCGGTGGGCAAGTCAATTTCTCTGACTATGCGCAAAAGT 4783
Oy 2877 gggggctgcgcactcgttctgcttggaaagaacaaatataagacagcgagta 2936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4784 ATCAGACGCCACATCATCATGATGATGAAAAAACCCCTTGTGATGACAGCTAAGCAC 4843
Oy 2937 agaagggaaaagcagacaacatgtgaaggaattagagtttcgagaagtccttctt 2996
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4844 GGAAGGCTTAATGCGCAACATGTGAAGGAATGTACATTTGGTGAAGTTGTATTCA 4903
Oy 2997 ctatccatgctgcagagatgtttcatctcctgcgtgtatataccctcagtatagcagag 3056
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4904 CTATGCCAACCCGACCGGACATCCAGTCTTCCAGGACTGAGCCGGAAGTAAAGAGG 4963
Oy 3057 aaagacagtagcatttggggagcagcgctgtggaagaaacacttctgttcaacttct 3116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4964 CCAAGACGCTGCTGTGTGGGAGCAAGTGTGTGGAAAGACAAAGTGTCACTCTCT 5023
Oy 3117 gcaagacttataagaccocgtgcaagacaagtctgttgaatgtgtgagtgatgcaaaaga 3176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5024 GGAGCGGTTTCTAGACCCCTTGGCAGGAAAGTGTGCTTGTATGGCAAAATTAAGCG 5083

```

```

Oy 3177 attgaatgacagtgctccgtccccaatagacaatgcttcccaagagccttgcctt 3236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5084 ACTGAATGTTCAAGTGGCTCCGAGCACACTGGGCAATCGTGTCGACGAGCCATCTGTT 5143
Oy 3237 caactgacagcatttgcctgaacatcgcctatgtgtgacacagcgtgtgtgtccattga 3296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5144 TGACTGACGACATGCTGGAACATTTGCCATATGAGACAAACGCGGTGTGTACAGGA 5203
Oy 3297 tgaatcaaaaagcgcgaatgacgaatatatccattcttatttgaagtctccgta 3356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5204 AGAGATCGTGAAGGCGAGCAAGAGGAGCCACATACATCTCTTCACTGAGTCACTCCTA 5263
Oy 3357 gaatacaacacacagttgacatgtaagagagcagagcttctgtgcgcgaacaaag 3416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5264 TAAATATGACATTAAGTAGAGACAAAGACACTAGCTCTGTGTGGCCAAACAAACG 5323
Oy 3417 actaactatgcaaggtctctcccaaaaacccaatttatttctgtgagatgagccac 3476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5324 CATTCCTATAGCTCGTCCCTGTTAGACAGCCCTCATATTTTGTGTGATGAGAGCCAC 5383
Oy 3477 ttcagcctcgataatgacagtgagaggtgtgtcaagatgccttgataaagcagagc 3536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5384 GTACCTCTGTGATACGAAAGTGAAGAGTGTGTCCAAAGAGCCCTGGACAAAGCCAGGA 5443
Oy 3537 gggaaagcagtcagtgagtgactcacagagctctcgaattcgaacgcagattgat 3596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5444 AGGCCGCACTGACATATGATGTTGCTACCGCTGTGCCACATCCGAATGCAAGATTAT 5503
Oy 3597 agtggtctgcacaatggaagataaaggaacacacacacacacacacacacacacac 3656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5504 AGTGTGTTTCAGATGACAGATGCAAGAGCATGCAAGAGCATGCAAGAGCTGCTGCACA 5563
Oy 3657 tcagacatatatttgaagtagtgaatgacagatcag 3694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5564 GAAAGCATCTATTTTCAATGCTCATGCTGCAGGCTG 5601

RESULT 4
5206352-3
; Patent No. 5206352
; APPLICANT: Roninson, Igor B.; Pastan, Ira H.; Gottesman,
; Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
; SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622, 836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892, 575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845, 610
; FILING DATE: 28-MAR-1986
; SEQ ID NO: 3
; LENGTH: 4669
5206352-3

Query Match 33.1%; Score 1223.6; DB 6; Length 4669;
Best Local Similarity 60.8%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 1269; Indels 132; Gaps 6;
Oy 162 caggttgaacctatattatgttgaataggtgtgtcgtccttgaatttggttacataca 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 760 cagtgatgcctattattacagtggaattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 819
Oy 222 gattcctgtgtatataactgacagcagacagcgaagagagatttgaacagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 820 gtttcatttgggtccgtgagcgtggaagacaataacaaattgaagaacagtttt 879
Oy 282 tcatcagtttggcagagacagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 880 tcatgtctataatgagcagagagatagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 939

```



OY	342	caatcgcatgaagaacatlgac---	aaatcgtgatggtatgtgaagataaagtctct	398
Db	940	caaccgacttaacagatlgatgctc	taagaattatacaagttatcttggacaataatlgaa	999
OY	399	gttgttcaaaacagatgcactctt	tcttcgaatggccgagatgtgtgttggtaaggcgc	458
Db	1000	gtctcttcagttcaatggcaacat	cttttcaacgtgttataagtaagattcaacagctgttg	1059
OY	459	gaaacctcaaccagtagtgaact	ctatccaagctctcttataatggtcttcagcggaacatg	518
Db	1060	gaagctaacacctgttgattttg	ccacacgcctcgtctcttgtagctgcaagctgtctctg	1119
OY	519	tctctagatggtgcacatcctat	tggaccaagtaagaattaaagtgcctatccaagcttgggc	578
Db	1120	ggcaaaagaaactatcctcatt	ctacgtataaagaaccttagcgtatggcaaaagctggagc	1179
OY	579	tgtggcagaagaagctcttct	tcaatccaaccgaagctacagctcttataaggcccaaggaa	638
Db	1180	agtagctgaagaggtctcttg	cgacgaattagaactgtgatcttcttggggacaanaaga	1239
OY	639	agaacttcaaaagctctctct	ttaataataacaagaatgtgtgtttatttcccca	698
Db	1240	agaac-----	-----	1244
OY	699	gtggtctactaagtgtgtctgt	ttnttgttaaggtatacaagaatctcaagaatgcaaa	758
Db	1245	-----	-----tgaagaagtacaacaaaactttagaagaagctaa	1278
OY	759	ggatttggcataaaagaagac	tataagcttcaaaagctctctgtgctgtgtgaactctt	818
Db	1279	aagaattgggataaagaaga	agctatacttaacagcaaatattctatagtggtgtcttctgtc	1338
OY	819	tatgaatgaacctatggaact	gtcttcttggtaaggaaacctctgtatcttaattcaatggaga	878
Db	1339	gatctatgcatcttattgctc	tctgctctgtgtatggacaaccttgcctctccagggga	1398
OY	879	acctgatatcatcatcggag	agctgttctgtgttctttagtgtaatccatatcgacgtga	938
Db	1399	-----atactatctatggc	aaatgaaccatctgtatcttcttccgtatataatgtgggcttc	1452
OY	939	tgtcatctggagcagcagctc	ccctcaacttgaaccttcgcacatgcccgaaggagctgcct	998
Db	1453	tagtgtgtgacaagcagctc	ccccaagcatgaagatttgcacaatgtgaaggagcagctta	1512
OY	999	tcatatttccaggtattatg	ataagaaccagtatagataaacttttccaagctgtgata	1058
Db	1513	tgaatcttccaagataattat	gataataagccaagtatlttgacaagctatcttcgaagaagtggca	1572
OY	1059	taaacctgatccatagaa	gaagcagctgtggaatttaaaatggtcttccatataccatc	1118
Db	1573	caaacacagataataata	agggaaatttgaattcagaagaatgtctcaacttccagttacccatc	1632
OY	1119	aagacacatcatcaagaat	ctctgaagctgaatctcagaatlaagtctggagagaacagt	1178
Db	1633	tcgaagaagaagttlaagat	ctcttgaaggccctgaacctgtgaagtgacaggtggcagaagct	1692
OY	1179	cgcctgtgtgcgtctcaat	ggcagtggtgggaagtaagtgtagtccagctcttcgaagaggt	1238
Db	1693	ggcctcgtgttggaaaca	gttgccgtgtgggaagacaacaacgctccagctgtatgcaagagct	1752
OY	1239	atatgtaccgagatgctct	tcaatctcgaatgggaatgtgaatgcagatccagagctttaa	1298
Db	1753	ctatgaccccaacagag	ggagatgtctgttctgtatggacaagatataatgaagccataaatgt	1812
OY	1299	gcgcgacatcatcgaa	ccataatgtggttagtcaagaagccctgtttgtctcggaacac	1358
Db	1813	aaggtttctaacgggaat	caatctgtgtgtgggttcgaagacccgtatgtttgtccaccac	1872
OY	1359	catcagtaacaataca	agaatgaagcagagatgtgactgcatgaagaagatgtgaagagc	1418
Db	1873	gatagtctgaaaaacat	ctgcgtatgtccgttgaagaatgtccacatgatatgaattgaagaagc	1932

QY	1419	agcaagggaagcaaatgvcgtatgattttatctatgagtttccatttaatttaatt	1478
Db	1933	tgctaaagaagccaattgcctatgactttatctatgaactgcctctaataattgacacct	1992
QY	1479	ggttagggaaagaagagcttcaaaatgagtgaggggacagaacaagaagatcgcaattgcctg	1538
Db	1993	ggtctggagagagaggggcccagcttgagtcgggtaggaagacgaagatcgccattgcacg	2052
QY	1539	tgcccttagcttgcgaaccccgaagatctcctcgtctgtagtaggcacgctgcagcttgacac	1598
Db	2053	tgccctggtctgcgaaccccgaagatctcctcgtctgtagtaggcacgctgcagcttgacac	2112
QY	1599	agaaagcaagtcagctcgtttcaagctcgcactgcgtggaagagcgcaaggttcgacttaaat	1658
Db	2113	agaaagcgaagcagcttggtttcaagctgcgtctcgtgataaaggccgaaaaaggttcggaccacat	2172
QY	1659	cgctgtagcacaacccgactttctactctctgaagtcgaagattgtagcttgagccccaagga	1718
Db	2173	ctgtatagcccatcctgcttgcctcagttcgttaatgctgcagctcactcgtgtgttcgattga	2232
QY	1719	tggaatctgctggcggagagaagagcagctgcagctgaactaahgycaaaacgaggtctatatla	1778
Db	2233	tgsgagctacttgtaggaagaagaatcatgatgactcatgtaagaagaaagcattactt	2292
QY	1779	ttcacctgtgtagtgcacagagatattaaagaagctgataacagatgtaggtctaaagacata	1838
Db	2293	caaacctgtctcaaatgtagcagacagaaagatgaagttcgtatgaaatlgacgtctgata	2352
QY	1839	cttcta-----ctgaaagaagaagccaactccttccttcgcactcgtg	1881
Db	2353	atccaaaggtgaatgtgatagcctcttggaatagcttctccaatgtatcaagatccagttcaat	2412
QY	1882	aagagcatca-----agtcgaagctcatgtaacagagcttgagga	1919
Db	2413	aagaaaaagatcaactcgttagaggtgctgcgtgtagtacaagccgaagcagaagaagcttag	2472
QY	1920	atccacccaatctaagaagataagctccctccgcgaagctcctctataaaatttaagtt	1979
Db	2473	taccaaagaagctctcgtgataagaagataaccacagtttcccttcttgagagattaaagagct	2532
QY	1980	aaacaagcctgaagggccttttggtttcctgggagacatggcttcgttcctaaatgtagaac	2039
Db	2533	aaatttaactgaaagggccttaattctgtctgtgtgattcttgcgcattataataatgtagag	2592
QY	2040	tgctatccagatatttccatctcactcttgccaaaatataaaccaatggtt---ggaaaataa	2096
Db	2593	ccctgaaccagcatttgcaataatatttccaagaatataaggggttttcaagaatgata	2652
QY	2097	tgataaaaacacatlaaagcagatgacagaatattatccatgatatctgcattgtgg	2156
Db	2653	tgatccctgaacaaacaaacagacagaatagtaactggtttcactatggtttctacagcccttg	2712
QY	2157	tgtaatttgcttttcagatttatctatcagacagagatatttttcgcgcagagcagggaat	2216
Db	2713	aattatttctttatcatatttctcctcaggggttccaatlttgccaaagcttgtagagat	2772
QY	2217	tttaacgattgatatgaacacttgcgcttcaagagccatgattatataacgatatgtgcctg	2276
Db	2773	cctaaccaagcgctccgataacatggttttcgatactcatgctcagacagagatgtaggtg	2832
QY	2277	gtttgataaaaagaaacagcacagagagcgttgacacaacatatgacatagatatagc	2336
Db	2833	ggttgatgacccctaanaaacacactgagatgactgattaccagagcttcgcacatgtagtgc	2892
QY	2337	acaaatcaagagacacaggttccagagattggcgtcttaacacaaaatgccaactaaat	2396
Db	2893	tcaagttlaaaggggtctaggttccagagcttgctgttaattaccacgaataatagcaaatct	2952
QY	2397	gggacttccagtttcaattccttcaataatgtagtagggagatgaaacctccatgactgag	2456
Db	2953	tgggacaggaataatataaccttcaactatggttggcacttaacacgcttactccttagc	3012
QY	2457	tatttctccagttacttgccgtgaagaaggaatgagattgaacccgacgaatgactggaattgc	2516

Db 3013 aattgtaccatctctgtcgaatagcagagatgtgtgaaatgaatgttgttggacaagc 3072  
OY 2517 caacaaagataagcaagaacttaagcagtgtgaaagatagcaactgaagcttggagaa 2576  
Db 3073 actgaagaataagaaagacagagtgctgggaagatcgtactgaagaatagaaaa 3132  
OY 2577 tatcgtactatagtgtcattaaacaagaagcccttgcagcaaatgtatgaaagat 2636  
Db 3133 ctccgaacccgttgtcttcttgactcagagcagaagttgaacatcgttctgcag 3192  
OY 2637 gcttcaactcaacaagaatcactcgaagaagcagaattttgaaagctgtatgc 2696  
Db 3193 ttgcaggttaccatacagaactcttggagaagcacacactcttggaaatcatttc 3252  
OY 2697 attcaagcagtcttataatatttgcctatgcagcagggttgcatttggagcattt 2756  
Db 3253 ctccaccaggaatgtatgtatttccctatgctgagatgttccgttggagcactt 3312  
OY 2757 aatcaagcttgacgaatgaccacagagggcatgttcatagttttactgcaattgca 2816  
Db 3313 gggtgcataactccttgagccttggagatgttctgttgaattttcagctgttct 3372  
OY 2817 ttggagctatggccatcgaaaaacgctcgttctgttgcctatcattccaaagccaatc 2876  
Db 3373 tgtgtccatgtgcgttgggagcagtcagtcattcttgcctcgtactatgcacaaagccaaat 3432  
OY 2877 gggtgtcgtgcagatctgttgccttgttggaaaagaacaataatagacagcgcagtc 2936  
Db 3433 atcagcagccacatcatcatgatcatgtgaaaaaaccccttggattgacagctacagac 3492  
OY 2937 agaaagggaagaaagccagacacatgttgaaggaatttgaagtttgcagaagctcttctt 2996  
Db 3493 ggaagggcctaaltgcgaacacatgtgaagaaatgtcatttggtagtgcatttcaa 3552  
OY 2997 ctatccatgtcccaagatgttccatccctcgcgtgcctacccccagctatgcagcagag 3056  
Db 3553 ctatccaccaccgagccagacatcccaagtgcttccaggaactgagcctggagggagaaagg 3612  
OY 3057 aaagacagtagcatttctggggagcagcgcgtgttggaaaaagcacttcttcaactct 3116  
Db 3613 ccagagcgtgtcgttgggagcagcagtgctgttgggaagcacaagtggtccagctcct 3672  
OY 3117 gcaagacatttatgaccccggtgcagaagcagtgctgttgaatgtgtgtgcagaaaga 3176  
Db 3673 ggaagcgttctacaccccttgcagggaggaagtgctgttgaatgtgcagaataaagcg 3732  
OY 3177 attgaatgtacagtggtcgtcttcccaatagcaatgttcccaagcgttgcctt 3236  
Db 3733 actgaatgttcaagtggtcgtccagcagccacttggcgtgttcccaagagcccatcctgt 3792  
OY 3237 caactgcagcatttctgagaacatcgcctatgtgtgacaaagcgtgtgtgtccattga 3296  
Db 3793 tgactgcagcatttctggaacatgtcctatgtgagaaacaagccgggtgtgtccagga 3852  
OY 3297 tgagatcaagaagcgcgaatgcagcaataatcattccttatttgaaggtctccctga 3356  
Db 3853 agagatctgtgagcagaagaaggccaacatcattccttcaatcagatcgtcctaa 3912  
OY 3357 gaaatacaacacacaagtgtgacactgaaagggacagccttctgcgcgcgcgaagaag 3416  
Db 3913 taataatgcactaaagttagagagacaagaactcagcttctgtgtgcggaacaag 3972  
OY 3417 actagctatgtcaaggggtcttctccaaaacccaataatttattgttggatgtgagccac 3476  
Db 3973 catgtccatagctcgttgcctttagagcagccatcatalatttgccttggatgaagccac 4032  
OY 3477 ttgagccctgcataatgcagtgtgagaagtggttccagcattgccttgcattaaagccagac 3536  
Db 4033 gtccagcttgcatacagaagaagtgtcttccaaagaagcccggtgcagaagccagaga 4092  
OY 3537 gggaagagcagctagtggtgcacatcagagcgtctcgtcaattcagaacgcagattgat 3596

Db 4093 aggcgcgaccctgcatgtgtatgttgcacgcgtgtccaccatccagatgcaagcttaat 4152  
OY 3597 agtgtgtcgcacacatgtgaagataaaggaagcaagcaatcatcaagctcctgagaaa 3656  
Db 4153 agtgtgttcaagatgtgcagagtcgaagcagcagcagcagcagcagcagcagcagc 4212  
OY 3657 tcgagacatatatttgaatgaatgtgagtcagtcagc 3694  
Db 4213 gaaagcattcatttccaatgtgtcagtgctccagcgtg 4250

RESULT 5  
US-08-583-276-18  
; Sequence 18, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:  
; APPLICANT: McDonagh, Kevin T.  
; APPLICANT: Nienhuis, Arthur  
; APPLICANT: Tolstoshev, Paul  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Carella, Byrne, Bain, Giffillan,  
; ADDRESS: Cecchi & Stewart  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: DNA V2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,276  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/332,444  
; FILING DATE: 31-OCT-1994  
; APPLICATION NUMBER: 07/887,712  
; FILING DATE: 22-MAY-1992  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4669 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: singular  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: Genomic DNA  
; US-08-583-276-18

Query Match 32.9%; Score 1217.2; DB 2; Length 4669;  
Best Local Similarity 60.7%; Pred. No. 0;  
Matches 2173; Conservative 0; Mismatches 1273; Indels 132; Gaps 6;

OY 162 caggttgaccctgtatattatgttgaatagtggtgtcgtcctgtatttctgttaccata 221  
Db 760 CAGGTATGCCCTATTATTTACAGTGCATTTGTCGGTGGCTGCTGCTTACATTCA 819  
OY 222 gattcctgtggaatataaactgcagcagcagcagcagcagcagcagcagcagcagcagc 281  
Db 820 GGTTCATTATTGTCCTCGCAGCGAGCAAGCAATTAACAAATTTAAGAAAAAGTTT 879  
OY 282 tcaatcagtttggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 341  
Db 880 TCATCTATTAATGCGACAGAGATAGCTGCTTGTGATGTGCACGATVGTGGCAGCTTA 939  
OY 342 cactgcagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 398







```
Db 2953 TGGAGCAGGAATATTAATCTTCATCTATGTTGGCACTAAGACCTGTACTCTTAC 3012
Qy 2457 tatgtctcagacttgcgcgtgacaggaatgattgaaccgcagcaatgactcgattgc 2516
Db 3013 AATTGTACCCATCATATGCAATAGCAGGAGTTGTGAAATGAAATGTTTGTGCGACAAAC 3072
Qy 2517 caacaaagaataagcaacttaagcaatgctggaagaatagcaactgaagcttggagaa 2576
Db 3073 ACTGAAAGATTAAGAAAGAACTGGAAGGTCTGGAAAGATCGCTACTGAAGCAATGAA 3132
Qy 2577 tataagctactagtcattacaaaggaagaaagccttcgagcaatgataagagat 2636
Db 3133 CTTCGGAACCGTTGTTTCTTTGACTCAGAGAGAAAGTTTGAACATATGATGTCTCAG 3192
Qy 2637 gcttcagactcaacagaataactctcgagaagaacacagatatttgaagctgtatgc 2696
Db 3193 TTTTGAGGTACCATACAGAACTCTTTGAGAAACACACATCTTTGGAATTAACATTTTC 3252
Qy 2697 attcagccatgccttatataatttcgcctatgcagcaaggttcgattgagacattt 2756
Db 3253 CTTCACCCAGCAATGATGATCTATTTTCCATCTGTGATCTTCCGTTTGGAGCTACTT 3312
Qy 2757 aatcaagctggaagaaatgaccccgagggcagtltcatalagtlttactgcaatgaca 2816
Db 3313 GGTGGCACATAAATCATCATGAGCTTGAGATGTCTGTAGATATTTTACGTGTGTCTT 3372
Qy 2817 tggagctatgagccatgcgaaaaagcgtgttggctcctgataattccaagccaatc 2876
Db 3373 TGGTCCCATGGCCGGGGGAGCAGTTCAATTTGCTCTGACTATGCCCCAAGCCAAAT 3432
Qy 2877 gggggctgcgacatcgtgttcgtgtgttggaagaacccaataatagagcgcgcgtca 2936
Db 3433 ATCAGCAGCCCATCATCATCATATGTAATAAACCCCTTGATTGTACAGCTACAGCAC 3492
Qy 2937 aagaaggaagaaagcagacacacatgtaaggaatttagatttcgagaagcttcttct 2996
Db 3493 GGAAGGCCATGATGCCGAACACATGTGAAGAAATGTCACATTTGGTGAAGTGTATTC 3552
Qy 2997 ctatccatgtgcgacagatglttcaatctcgtgtgttataccctcagatattgagcgag 3056
Db 3553 CTATCCCAACCCACCGGAGATCCAGTGTCTTCAGGAGACGACCTTGAGAGTAAGAGG 3612
Qy 3057 aagaagcagtagatttggggagcagcgctgtggaagaacacttcgttcaacttct 3116
Db 3613 CCAAGACGCTGCGCTCTGtgggGCAAGtGCTGtggGAGACACAGtGtCCAGCTCT 3672
Qy 3117 gcaagacttataagccccgtgcaaggaagtgctgttgaatgtgtgtagatgacaaga 3176
Db 3673 GAGAGGGTCTACGACACCTTGGCAGGAAAGTGTCTGTATGGCAAGAAATAAAGCG 3732
Qy 3177 attgaatgaaagtgctcgttcccaatagcaatcgttctcgaagcgtgtgtctt 3236
Db 3733 ACTGAATGTCTAGTGGCTCCGAGACACCTGGGACATCGTCCAGGACCCATCTGTT 3792
Qy 3237 caactgcaacatgttgagaacatcgctatgtgtaacaacgcgtgtgtgctcattaga 3296
Db 3793 TGACTGCAACATGTCTGAGAAACATTCCTATGAGACAAACACGCGGTGTGTACAGGA 3852
Qy 3297 tgaagatcaagaagcgcgaatgacgaataatcaatcttattatgaagcttccctga 3356
Db 3853 AGAGATCGGAGGAGCAAAAGAGGCCAACTACATGCTTCATGAGTCAAGTCAAGCTTA 3912
Qy 3357 gaaataacaacacagatgtgagctgaaggaagacagcttcttcgagcgcaaaagaag 3416
Db 3913 TAAATATACACTAAAGTGTGAGACAAAGAACTCAGCTCTGTGGGCGCAAAAAGG 3972
Qy 3417 actgactatgcaagcgctcttccaagaaccgaatattatgttgaatgagcgac 3476
Db 3973 CATTTGCATAGCTCGTGGCTTTGTAGACAGCTCATATTTTGTGATGAAGCCAC 4032
Qy 3477 ttcagcctcgaataatgacagtgaaggtgtgtcagcatgsccttgataaagcagagc 3536
Db 4033 GTCAGCTCTGTATACAGAAAGTGAAGAGGTGTCCAAAGGCCCTTGAGCAAAAGCAGAG 4092
```

```
Qy 3537 gggaaagcactgcctagtggtcactcaacaggtctctgcaattcagaagcagattgat 3596
Db 4093 AGCCCGCACCTCATTTGATTTGCTACCGCTGTCCACACCAAGATGACAGACTTAAT 4152
Qy 3597 atgtgtctgcacaaatggaagaataaaggaacgaactcaagaagctctcctgagaa 3656
Db 4153 AGTGGTGTTCAGATGCGAGAGTCAAGAGCATGGCAGCATGACAGCTCTGTGGACA 4212
Qy 3657 tcgagacataatlttaagtlagatgcaagtcag 3694
Db 4213 GAAGGACATCATTTTTCATATGTGCTAGTGTCCAGGCTG 4250

RESULT 7
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: SIKIC, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg. No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-784-649A-1

Query Match 32.8%; Score 1214.2; DB 2; Length 4264;
Best Local Similarity 60.5%; Pred. No. 0;
Matches 2163; Conservative 0; Mismatches 1283; Indels 129; Gaps 5;
```



OY	342	caacgcgatgaacagacttgaacaaatcgtgatgtatcttggaataagatgtcctgtt	401
Db	657	CACCGGACTTTACAGATGATGTCTCCAAAGATTAAATGAAAGAAATTTGGTGCACAAATTTGGAAAT	716
OY	402	gttccaaacatgtctcactcttcttcgatctggccgagctgtgttggtgaaagcgctgaa	461
Db	717	GTTCACGATGCAATGGCACAACATTTTTCACCTGGGTTTATAGTAGATTTTACACGGTGTGGAA	776
OY	462	actcaacctagtgactctatccacgtctccctcttaataatgctcagcgagcaatgttc	521
Db	777	GCTAACCCCTGTGATTTTTGGCCATCAGCTCCTGTTCTTGGACGTGCAGCTGCTGTGGCC	836
OY	522	taggatgtctatccatctatggcccgtaagaatlaagtgctctatccaaagctgggctgt	581
Db	837	AAAGATCTATCTTCATTTTCTGATTAAGAACCTTTAGGGTTGGAAAAAGCTGGAGCAGT	896
OY	582	ggcagaagaagctctgtgacaaatccgaacagctacatagccttlaagcccgagaaaga	641
Db	897	AGCTGAAGAGGCTCTTGGCAGACATTTAGAACGTGATTTGCAATTTGGAGGACAAAGAAAGA	956
OY	642	acttcaaaagylcttccctttaaataaacaagataltgcttgttltatlttccccagtg	701
Db	957	AC-----	958
OY	702	gctactaatgtgtctgtcttntcttgaagatatacagaatctccaagaatgtcaaga	761
Db	959	-----TTGAAAGGTTCACCAACAAATTTTGAAGAAAGCTTAAAG	995
OY	762	tttggcataaagaagactatagcttcaaaagtgctctctgtgtgtgtactcttatt	821
Db	996	AATTGGGATTAAGAAAGCTTTTACAGCCAAATTTTCTATAGGTGGCTTCTCTCTAT	1055
OY	822	gaatggaaacctatggaacttcttcttggtaatggaacctcttgatctcttaatlygaaac	881
Db	1056	CTATGCACTTTATCTCTGGCCTTTCGGTATGGAGACACCTTGTCCTCTCAGGGGA---	1112
OY	882	tgatataccaatcggagacttctctgtcttcttctttagtgtaatccatagcagtiatgt	941
Db	1113	---ATATTTCTATTGGAACAACTACTCACTGATCTTTCTGTATTAAATTTGGGGCTTTTAA	1169
OY	942	catggagcagcagtgctcctcaacttgaaccttgaaccttgcgaatagccgagagctgtcttca	1001
Db	1170	TGTTGGACAGGCATCTCCAAAGCATGTAGCAATTTGCAATGCAAGAGGACGCTTATGA	1229
OY	1002	tatttccaggttctatgataagaaccagactataga taaacttctccaagctgagataaa	1061
Db	1230	AATCTCAAGTATTATGTATTAATAGCCAGTATTGACACTTTTGGAAGAGTGGGCCAAG	1289
OY	1062	acctgaatccaatagaagaactgttgaatttaaataatgttcttccaatlatccatcaag	1121
Db	1290	ACCGATATATTTATTTAGGGAATTTGGAATTTCAAAATGTTCACTTCAGTTACCCATCTCG	1349
OY	1122	accatctatacaagttcttgaaaggtcttgaaatctcagaatlaagctctgagagacagtgcc	1181
Db	1350	AAAAGAACTTTAAGATCTTTGAAGGCGCTCGAACCTCGAAGGTGCAGAGTGGGCAACGCTGTGC	1409
OY	1182	cttgctgcttccaatgtgcagcttggaagaagtaacgtagtccaagctcttgcgaaggttata	1241
Db	1410	CTGTGGTGAAGAACGTGTGCTGTGGGAAGACAAACAGTCCAGCTGATGTCCAGAGGCTTA	1469
OY	1242	tgatccggaatgagtcttatacaltgvtgtaatgaaatgaatcaagacttaaatatgtcg	1301
Db	1470	TGACCCCAAGAGGGAGTGTGCTAGTGTGATGGACAGATTTTGAACCATTAATATGTAAAG	1529
OY	1302	gcattatccgagacatatattggaatgtgttagtcaagaagcctgttctgtccggagccacat	1361
Db	1530	GTTTCTACGGGAATTCATTTGGTGTGTGATCGTAGCAACCTGATTTGTTGCCACACCAT	1589
OY	1362	cagtaaacataatcaagataggaagagatgagtgaactgttgaaagagatggaagaagcgcc	1421
Db	1590	AGCTGAANAACTTTGGCTATATGGCCGTGTAATAATGTCCACCATGATGTAGATTTGAANAAGCTGT	1649
OY	1422	aaggaagcaaatatcgatgatatttataatcaltgagatcttccataaatttaataatgtgt	1481

Db	1650	CAGAAGAGCCAAATGCTCTATACCTTTTTCATGAAACTGCTCTCAATAATTGGACCCCTGCT	1709
QY	1482	agggagaaagagactcnaatlgagtgagagggcagaacaagagatcgcgaattgcgtctg	1541
Db	1710	TGGAGAGAGAGGGGCCCAAGTTGAGTGGTGGCGAAGACGAGAGATGCGCATTCGACGTC	1769
QY	1542	cttagtctcgaaaceccaagaattctgatattagatgagctacgctctccctgfatcaga	1601
Db	1770	CCCTGGTTGCGAACCCCAAGATCCCTCGTGGTGGAGGCCAGCTGATGCGCTTGACACGA	1829
QY	1602	aagaagatcagctcgttcaaacctcgaactgagagagggcagaagtgagactacaaatcgt	1661
Db	1830	AAGGAAAGCAGTGGTTAGGTGGCTCTGGATMAAGGCCAATAAAGGTCGACACCATTGCT	1889
QY	1662	gttagcacacggaattctctctactctatctgaagtgagaaattgattgtgacccaaagaatg	1721
Db	1890	GATAGCTCATCGTTTGTGCTACACTTGCTGATAGCTGATGACGTACGTGCGTTGATGATGG	1949
QY	1722	aatgcctggcggagagaagagacacatgcctgaactaatgacgaacgaggtctatatattc	1781
Db	1950	AGTCATTGTGGAGAAAGAAATCATCATATCAATCAGTAAGGAAGAAAGCATTTACTTCAA	2009
QY	1782	actgtgatgtcgaagataattaaaaagctgatgacagatgaggtcaatgacatctc	1841
Db	2010	ACTGTGCACAAATGACGACGACGAGAAATGAAGTGATTTAGAAATTCGACGTGAATTC	2069
QY	1842	ta-----ctgaagaagaagcaaacatctctctcgtgaactctgtag	1884
Db	2070	CMAAAGTGAAATTCATGCCCTTGGAAATGTCCTTCAATGATTCAAATCCTCAGCTAATATAG	2129
QY	1885	agcalca-----agtcagactcattgacaaagcctgagaaatc	1922
Db	2130	AAAAAGATCAACTGCTGAGAGTGCCTGGATGCACAAAGCCCAAGACAGAAAGCTTAGAC	2189
QY	1923	cacccaatctaaagagataagctctccctgaagctctctcataataaatlttaagttaa	1982
Db	2190	CAAAAGAGCTCTGGATCAAAAGTATACCTCCAGTTTCCCTTTGGAGATTATGAAACCTAAA	2249
QY	1983	caagcctgaagtgagccttgtagttcttgaggacaattgagctctgtctcnaatgagaaatc	2042
Db	2250	TTTAACTGATGAGCTTATTTTGTGTGTGTGATTTTGTGCANTTAATAATGCAAGCGCT	2309
QY	2043	tcatacgaattcttccatactcttgcnaaataataaacaatgltt---ggaanaatgaa	2099
Db	2310	GCAACCAAGCATTTGCATTAATATTTTCAAAAGATATAGGGTTTTTACAAAGATGATGAA	2369
QY	2100	taaaacacacattaaagcattgacagaaattatccatgatattcgtcaatttgggtgtc	2159
Db	2370	TCTGAAACCAAAAGACAGAAATAGTAACCTGTGTTTACATTTGTTTCAGCCCTTGAAAT	2429
QY	2160	tattgtcttgtagattatcttcatctgaaggaattattttagcagagcagggagaaattt	2219
Db	2430	TATTTCTTTTATTAACATTTTTCCTTTCAGGGTTTTCACATTTGGCAAAAGCTGGAGAGATCT	2489
QY	2220	aacgatgagatlaagacaacttgccttccaaagccatgctatata tcaggaatatgacctggt	2279
Db	2490	CACCAAGGGGCTCCGATACATATGGTTTTTCCGATCCATGCTCAGACAGATATGATGGTT	2549
QY	2280	tgatgaaaagaaaacagacagagaggtcttgacaacaatatatagccatagatatagaca	2339
Db	2550	TGATGACCTCAAAAACACCACTGAGACATTGATACCAAGCTCCGCAATGATGCTGCTCA	2609
QY	2340	aattcagagagcaacaggtctcgaggttggcgtcttaacaaanaatgaaactaacatagg	2399
Db	2610	AGTTAAAGGGCTATATAGTTTCCAGGCTTCTGTATTTACCCAGATATATGCAAAATCTTGG	2669
QY	2400	acttcaagttatcttcccttatacatgtagtggagagaaatccctgattctgagat	2459
Db	2670	GACGAGATATATTTATCTCTTCACTATGTGTTGGCAACTAACAACGTTCATCTTAGCAAT	2729
QY	2460	tgctccagttactgctgtagcagaatgtatgnaaccgcagcaatgactggaatttgcaa	2519

```

Db 2730 TGTACCCATCATTCGAATAGCAGAGGTTGTTGAATGAAATGTTGTCGACAAAGCACT 2789
OY 2520 caaagataagcaagaacttaagcatgcttgaagaatagcaactyaagctttgagaatat 2579
Db 2790 GAAAGATAAGAAAGAAACACTAGAGAGGTGCTGGAGAGATCCCTACTCAGAGCAAGAAAGATT 2849
OY 2580 acgtactatagtgcatataaagaagaagccttcgaagaatgtatgaagaagatgct 2639
Db 2850 CCGAACCGTTTCTTCTTGATGCTCAGAGACAGAACTTGACATATGTATGCTCAGAGTTT 2909
OY 2640 tcgaactcaacaagaataaccccgaaagaacagaattatlggaagctgtatgcat 2699
Db 2910 GCAGGTACCATACAGAAACCTTTGAGGAAGACACACTTTTGGAATTTACATTTTCTT 2969
OY 2700 cagcgaagccttatatatatttgcttaagcagcagggtttcgaatttgagccttaata 2759
Db 2970 CACCAGGCAATGATGATTTTCTTATGCTGATGATTTCCGATTTGGAGCTCTAGTGGT 3029
OY 2760 tcaagctggaagaaatgacccagagagcatgttcatagtttttaactgcaattgatalgt 2819
Db 3030 GGCACATTAACATCATGACCTTTGAGAGATGCTGTGTAGTATTTTACAGCTGTGCTTTGG 3089
OY 2820 agctatggcatalcggaaacacgctcgttttggctcctgaatalcccaagcacaatcggg 2879
Db 3090 TGCCATGGCCGTGGGGCAAGTCAGTTTATTTGCTCTGACTATGCAAGCAAGCAAAATATC 3149
OY 2880 ggtcgcgcatctgttgccttggttggaaagaaccaaataagacagcgcagtcaga 2939
Db 3150 AGCAGCCCACTCATCATGATCATGTAAGAAAAACCCCTTTGATTGACACTCAGACACGGA 3209
OY 2940 agggaaaaaagcagacagacatgtgaagaagaattagagtttgaagaagcttcttcttca 2999
Db 3210 AGGCTTAATGCCGAGACACATTTGAGAGAAATGTACATTTTGTAAGTTGTAATTTCACTA 3269
OY 3000 tccatgtccccaagatgtttcatcctccgttgcgttcaatccctcagttatlgacgagaa 3059
Db 3270 TCCACCCGACCGGACATCCAGTGCCTTCAAGGACTGAGCTGGAGGTGAAGAGGCGCA 3329
OY 3060 gacagtagcatctgttgggagcagcgcgtgttggaaagaacacttcttcaacttgcga 3119
Db 3330 GACCTGCTCTGTGGGACACAGTGCCTGTGGGAAGACACAGTGCACAGCTCCTTGA 3389
OY 3120 gagaacttaagcccggtgaagacagatgctgttgaatgtgtggaatgcaagaatt 3179
Db 3390 GCGGTTTACGACACCCCTTTGGAGGAAAGTGTCTGTATGCAAAAGAAATTAACGCACT 3449
OY 3180 gaattacagtggtcgttcccaaatagcaatcgttcccaagagcctgtgtcttcaa 3239
Db 3450 GAATGTTCAGTGGCTCCGAGACACCTGGGATGCTGCCAGAGCCCATTCCTTTGA 3509
OY 3240 ctgagacattctgaagaacatcgcctatagtgaacaacagccgtgtgtgccaattagatga 3299
Db 3510 CTGCGACATTCCTGAGAACATTTGCTATGAGACAAACGCGGTGGTGTCAACAGGAAGA 3569
OY 3300 gatcaagaagccgaaatgcaagaatcatcatcttatttgaagcttccctcgagaa 3359
Db 3570 GATTGTGAGGGCAGAAAGAGGCGCAACATATATGCTTTACATGATCAGTCCCTTAATTA 3629
OY 3360 atacaacacacaaglttgactgaagaagcagcattcttgcggcgacagaacaagaact 3419
Db 3630 ATATAGCACTTAAGTAGAGAGACAAAGAACTCAGCTCTTGCTGGTGCAGAAACACGAT 3689
OY 3420 agctatgtcaagggctcttcccaaaaaccaaatttatttattgttgaatgagccacttc 3479
Db 3690 TGCCATACCTCGTGCCCTTTGTTAGACACCTCATATTTTGTGTTGATGAAGACACATC 3749
OY 3480 agccctcatagacagctgaagaagltgttcaagcatgccttgaataagcagaagcg 3539
Db 3790 AGCTTGATATACAGAAATGAAAGGTTGTCCAGAAAGCCCTTGAGCAAGCAAGGAAG 3809
OY 3540 aaggaacatgcbtagtgcactacacaggtcctctgcaattcaagaacgagatttgaatg 3599
Db 3810 CCGCACCTGCAATTTGATTTGCTCACCGCTGCTCCACATTCAGATTCAGACTTAATAGT 3869

```

```

OY 3600 gttctgcacacatggaaagataaaggaaaggaactcatcaagagctctctgagaatcgc 3659
Db 3870 GGTGTTTCAGATATGCAAGTCAAGAGACATGCGCATGACGATGACGATGCTGGCACAGAA 3929
OY 3660 agacatatatttaagtagtgaatgcacagtcag 3694
Db 3930 AGGCATCATTTTTCATATGATGTCAGTGTCCAGGCTG 3964

RESULT
8
US-08-784-649A-5
: Sequence 5, Application US/08784649A
: Patent No. 5830697
: GENERAL INFORMATION:
: APPLICANT: SIKIC, Branimir I
: APPLICANT: Chen, Gang
: TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
: TITLE OF INVENTION: CYCLOSPORIN MODULATION
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 2200 Sand Hill Road
: CITY: Menlo Park
: STATE: CA
: COUNTRY: USA
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/784,649A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Sherwood, Pamela J
: REGISTRATION NUMBER: Reg. No. 5830697 36,677
: REFERENCE/DOCKET NUMBER: 06037/007001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-322-5070
: TELEFAX: 415-854-0875
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4264 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-784-649A-5

Query Match 32.8%; Score 1212.6; DB 2; Length 4264;
Best Local Similarity 60.5%; Pred. No. 0;
Matches 2162; Conservative 0; Mismatches 1284; Indels 129; Gaps 5;

```





Db 2790 GAAAGATAGAAAGAACTAGAAAGTGCTGGAGAGATCGCTACTGAGCAATAGAAACCTT 2849  
Oy 2580 acgtactatagtgatcatttaacaaaggaagcccttgagaaagtatgaagatgct 2639  
Db 2850 CCGAACCGTTTCTTTTCTTGTACTCAGGAGCAGAGATTGTAACATATATGCTCAGAGTTT 2909  
Oy 2640 tcagactcaacacgaataacacacgaagaaagacagatatttggaagctgtatgact 2699  
Db 2910 GCAGGTACCATACAGAAACCTTTGAGAGAACACATCTTTGGAATTTACATTTCTT 2969  
Oy 2700 cagcaatgcttataatatttgctatgacagaggttctgatttgagcttataat 2759  
Db 2970 CACCCAGGCAATGATGTATTTTCTTATGATGTATTCGGTTTGGAGCTACTTGGT 3029  
Oy 2760 tcaagctggaagatgaccacagagggatgttcaatgattttactgaattgctatg 2819  
Db 3030 GGCACATTAACATCATGACCTTTGAGAGATGCTGTATATTTTTCAGCTGTGCTTTGG 3089  
Oy 2820 agctatgacacacgaagaaagctgcttctgctcgaatattcgaagcccaatcg 2879  
Db 3090 TGCCATGGCCCTGGGCGAAGTCATGTCTTGTCTCTACATATGCAAAAGCCAAATATC 3149  
Oy 2880 ggtctgacatctgttctgttctgttgaagaaacaaataatagaacagccgactca 2939  
Db 3150 AGCAGCCCATCATCATCATGATGAAACCCCTTTGATTTGACAGCTACAGACGGA 3209  
Oy 2940 agggaaaaagccagacacatgtgaaggaatttagatttgcgaagctcttcttct 2999  
Db 3210 AGGGCTTAATGCGCAACATTTGGAAGAAATGTCATTTGGTGAATTTGATTAACAT 3269  
Oy 3000 tccatgtgccacatgttcttccatcctcgtgcttccatccctcaatgattgacgag 3059  
Db 3270 TCCACCGGACCGGACATCCAGTCTTCAGGAGCTGAGCTGGAGGTGAAAGAGGCA 3329  
Oy 3060 gacagtagcatttgbgagagcagcgtgtgagaaagcaacttcttcaacttgc 3119  
Db 3330 GACCTGTGCTGTGTGGGACAGAGTGTGGAAAGACAGAGTGTCCAGCTCTGGA 3389  
Oy 3120 gagacttatacccccgtgacagcaagtgctgttgaatgtgtgtgagtgcaaa 3179  
Db 3390 GCGGTTCTACACACCCCTTGGAGGGAAGTCTGCTTATGGAACAAATTAACCGACT 3449  
Oy 3180 gaattgaagctgctcgttcccaatagcaatgcttcccaagccctgtgcttcaa 3239  
Db 3450 GAATGTTAGTGCTCCAGACACACCTGGCATCTGTGCCAGAGGCCCATCTTTTGA 3509  
Oy 3240 ctgacgacatgctgagaaacatgcctatgtgtgacacacagccgtgtgtgcaat 3299  
Db 3510 CTGACGATTTGCTGAGACATTTGCTATGAGACAAAGCCGGGTGTGTCACAGAGA 3569  
Oy 3300 gatcaagaagccgcaaatgacgaataatcatttatttgaaggtctccctgaa 3359  
Db 3570 GATTGTGAGGCGAGAAAGAGGCGCCACATCATCTTCATCGAGTCACTCCATTA 3629  
Oy 3360 atacaacacacaagtgtgacgaagagcaagcttctgtgacgagccgaagaa 3419  
Db 3630 ATATATGCTAAAGTAGAGACAAAGAACTGAGCTCTGTGTGGCCGAGAAACAAGCAT 3689  
Oy 3420 agctattgcaagggctcttcccaaaaacccaataatttattgtgtgagtgaacact 3479  
Db 3690 TGCCATACCTGTGCGCCCTTTTGAACACGCTCATATTTTGTGTTGATGAAAGCAG 3749  
Oy 3480 agccctcgataatgaagtgaggaagtggttcaagcatgccccttgaataagcag 3539  
Db 3750 AGCTCTGATATCAGAAAGTGAAGGTGTCCAAAGAACCCGAGCAAGCAAGAGG 3809  
Oy 3540 aaggaacagctagtgatcattcaagagctctctgcaatlttaaggaagcaggtt 3599  
Db 3810 CCGACCTGCTATTGTGATTTGCTCACCGCTGTCCACATCCAGAAATGCAAGATTAT 3869  
Oy 3600 gttctgacacatggaagataaagaaacaggaactcatcaagagcccttgagaatcg 3659  
Db 3870 GGTGTTTCAAGATGCGACAGTCAAGGACATGGACGATGACGAGCTGTGCGACAGAA 3929

Oy 3660 agacatatatttaagttgaatgcacagcag 3694  
Db 3930 AGCATCTATTTTTCAATGCTCAGTGTCCAGGCTG 3964

## RESULT 9

US-09-120-513-1  
; Sequence 1, Application US/09120513

; Patent No. 6023160

; GENERAL INFORMATION:

; APPLICANT: Brun, Kimberly

; APPLICANT: Chenev, Richard

; APPLICANT: Ellens, Harma

; APPLICANT: Field, John

; APPLICANT: Yue, Lin

; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE

; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY:

; ZIP: 19406

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/120, 513

; FILING DATE: 22-JUL-1998

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: King, William T

; REGISTRATION NUMBER: 30,954

; REFERENCE/DOCKET NUMBER: GP50008

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-270-5015

; TELEFAX: 610-270-5090

; TELETYPE:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4233 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Genomic DNA

US-09-120-513-1

Query Match 30.3%; Score 1121; DB 3; Length 4233;  
Best Local Similarity 59.0%; Pred. No. 0;

Matches 2114; Conservative 0; Mismatches 1340; Indels 129; Gaps 6;

Oy 162 caggttgacccctgatttattgttgaatagtgctgtgctgctgatttgggtacataca 221  
Db 355 CATGACGCGCTACTTATTACAGCGGCGATTTGTCGCGTGTGCTCATGTTGCTTACATCA 414  
Oy 222 gatttccttgtgtatataactgacagcagacagacacaaagagattcgaataacagtttt 281  
Db 415 GGTTCACATTTGTGCTCCGCGAGCTGGGAGACAAATACACAAGATTAGCAGAACTTTT 474  
Oy 282 tcaatcagtttggacacaggaatcagctggtttatagctgtgacatcgtggaactaa 341  
Db 475 CCATCCCATCATGATTAATTCAGAGATAGCGTGTGTTGACGTAATGACGCTGGGGACTCA 534  
Oy 342 cactgcatgaca---gacatgacaaaatcagtgatgatttggagataagattgctct 398  
Db 535 CACCGGCGCTCACAGATGACGCTCTCCAAATTAATGACGGAATTGGTGACAAACTTGAAT 594

[illegible]

Db 2668 GAAGGACAAGAAAGAGCTAGATCTCGGAGAGATCGCTACAGAAACATTTGAAACCTT 2727  
QY 2580 agtactataagtgatcaatacaaggaaagccctcgagcaaatgataagagatgct 2639  
Db 2728 CCGCAGCTGTGTTCTCTTTGACTAGGAGAGCAAGTTTGAACATATGATGAGCCAGAGCTT 2787  
QY 2640 tcaagctcaacaagaataactctgaagaagcacagatattgaagtgatgacatt 2699  
Db 2788 GCAGATACCATTCACAGAAATGCTTTGAGAGAACACACGCTTTGGATACCTTGCCCTT 2847  
QY 2700 cagccatgcttataatatttgcctatgcagcagggtttcgacttgagccattat 2759  
Db 2848 CACCCAGGCGCATGATTTATTTTCCATGCTCTGTTCCGGTTCGGCTACTTGCT 2907  
QY 2760 tcaagctgaagaatgaagccagaggacatgcttcaagtttactgcaattgacatag 2819  
Db 2908 GGCAGAGAACTCAGACCTTGAAGAAATGTATGTGATTTTCTGCTGTGCTTTG 2967  
QY 2820 agctatgccaatcggaagaaagctcgcttgccttgatatttccaagccaaatcg 2879  
Db 2968 TCCCATGCGAGCAGGAAGAAATCCAGTTTCAATTCCTGACTAGCGGAAGCCAAAGCTTC 3027  
QY 2880 ggcctgcacatgcttgccttgcgaaagaacaaatataagacgcgagtcaga 2939  
Db 3028 AGCATTCACCATCATCAGATCATGATGAGAAATCCCGGATTTGACAGCTACAGCAGCA 3087  
QY 2940 agggaaaaagccagacacatgtaaggaattagaatttcgaagaagctcttcttcta 2999  
Db 3088 GGGCTTTGAAGCTTAATTTGGTTGAGAAAGTAATGTAATTAATGATGATTTCAACTA 3147  
QY 3000 tccatgctgcagagatglttctatctcgttgcttaccctcagatgagcagga 3059  
Db 3148 TCCCAACCCGACCAACATCCAGTCTTCAGGAGACTGAGCTTCGAGCTGAAGAGGGCA 3207  
QY 3060 gacagtaagcattgtggaagcagcgcgtgtggaagaaacacttctgttcaacttgcga 3119  
Db 3208 AACGCTTCGCCCTGCTGGTGGGAGCAGTGGCTGGGGAAGATACAGTGTCTCAGCTGTCA 3267  
QY 3120 gagacttatagaccccgctgcaagcaagtgctgttgaatgctgtgaatgaagaat 3179  
Db 3268 GCGCTTCTTCAACCCCATGCTGTAACATGTTTCTAGATGCAAGAAATTAACATCACT 3327  
QY 3180 gaatgaagtgctgccttcccaatagcaatgcttcttcaagccgtgtcttctca 3239  
Db 3328 CAACGTCACAGTGGCTCCGCGCCACCTGGGCAATGTGTCCAGGAGCCCATCTGTTTGA 3387  
QY 3240 ctgcagcattgtggaacatgcctatgtgtaacaacgcgtgtgtgcatatga 3299  
Db 3388 CTGCAAGCATCACGAGAAATCGCTACGAGAGACAACACCGTGTCTCATGAGGA 3447  
QY 3300 gatcaagaagccgaatagcagaataatcattcttattgaagtgcttccctgaaga 3359  
Db 3448 GATTCGTGAGGGCGCCAGAGGAGCAACATCCACAGTTCATTCGACTCTGAGAA 3507  
QY 3360 atacaacacacaagtgtgactgtaaggaagcagcttctgcgcgcaaaacaaagact 3419  
Db 3508 ATACAAACACAGAGTGGGAGACAAAGGACCTAGCTGTGGGCGGCAAGACAGGCAT 3567  
QY 3420 agctatgcaagagctcttctcaaaaacccaaatattatgttgaatgaagccattc 3479  
Db 3568 GCGCATTCGGCGCGCTGTGTACAGACAGCTTCATCTTTCGTGATGAGAACGACATC 3627  
QY 3480 agcctcgaataatgaatgagaagtggttcaagcattgccttgataaagcaagcagg 3539  
Db 3628 AACTCTGTGATGAGAGAGTGAAGAGTCTCCAGAGAGCGCTGAGCAAGAGCAGGAAG 3687  
QY 3540 aaggacatgctcctgaatgtaactcaagagctcttgaattggaagcagattgata 3599  
Db 3688 CCGCACTCTCATTTGTGATGCGCACCGCTGTCCACATTCAGAAAGCAGACTGTGCT 3747  
QY 3600 ggtctgcacaaatggaataaagaagaagcaagcaatcacaagactcctgaagactg 3659

Db 3748 GGTGATTTCAGAAAGCCAGAGTCAGAGCAGCGACACCCACAGACAGCTGTGCGCAGAA 3807  
QY 3660 agacataatatttaagttagtgaatgcagagtcagtgacgtga 3702  
Db 3808 AGGCATCTATTCTTCGATGATGCTGAGGCTGAGAGCAAGAGCGCTCA 3850  
RESULT 10  
US-09-450-105-1  
; Sequence 1, Application US/09450105  
; Patent No. 6169166  
; GENERAL INFORMATION:  
; APPLICANT: Kimberly Anne Brun  
; APPLICANT: Richard James Chenery  
; APPLICANT: Harma Eilens  
; APPLICANT: John Anthony Field  
; APPLICANT: Lin Yue  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES  
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF  
; FILE REFERENCE: GP-50008-D1  
; CURRENT APPLICATION NUMBER: US/09/450,105  
; CURRENT FILING DATE: 1999-11-29  
; EARLIER APPLICATION NUMBER: 09/120,513  
; EARLIER FILING DATE: 1998-07-22  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 4233  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
US-09-450-105-1

Query Match 30.3%; Score 1121; DB 4; Length 4233;  
Best Local Similarity 59.0%; Pred. No. 0;  
Matches 2114; Conservative 0; Mismatches 1340; Indels 129; Gaps 6;

QY 162 caggttgacccgtatattatgtgaatagtgctgtccttgatatttgyttaactaa 221  
Db 355 catgtacgctactataacagcgatgtgcggtgtgtcatatgtgctacatca 414  
QY 222 gatctcctgtgattataactgcagcagcagacagagagattcgaaacagtttt 281  
Db 415 ggttctacttgytgcctgcagcgtggaagacaataacaaagattagcagaagtttt 474  
QY 282 tcaatcagtttgcagcagacatcgcgtgtgtgaatgctgtgacatcgttgaactaa 341  
Db 475 ccatgcacatgaatcagagatagcgtgtgtgacgtgaatgacgcgtggagctcaa 534  
QY 342 cactgcagatgaca---gaactgaacaaatcagtgatgatatggaagataagatgctct 398  
Db 535 caccgcgctcagcagatgagctctccaaataatgaacgaagatgtgtgacaacttgaat 594  
QY 399 gttgttcaaaaatgtctacttcttgcatttgcattggtgcgtgcaggtgtgttgaaggctc 458  
Db 595 gttcttccatgcataacagacatttccagccggttttataatgagttataatagtgctg 654  
QY 459 gaaactcaaccctagtgactctatccacgctctccttataatgcttccagcggcagcatg 518  
Db 655 gaagctaacccctgttaatttgcgtcgaagccctctattgtgttcatcctgcatgtg 714  
QY 519 tcttagatggtcatcattatgacccagtaaggaatgaatgctatccaaagctggtgc 578  
Db 715 ggcnaaggtactgactctatccataaagaaacccaggttactgcaaaagcttggagc 774  
QY 579 tgtgagaagaagcttctcatcaatccgaacagtcataagcctttagggccagaagaa 638  
Db 775 agttgcgaagaagcttctgcagccatccgaagctgtgattgcttggggacaaaagaa 834  
QY 639 agaacttcaaaagcttcttccatttaataataaacaagatatgcttgtttatttccca 698  
Db 835 ggaac----- 839

QY 699 gtcgactaagtgtgtctgttnttltgaaagatatacagaaatcacaagtacaa 758  
Db 840 -----ttgaaagatacaataaatttgaagaagctaa 873  
QY 759 ggaatttgcataaaagaactataagcttcaaaagtctctgttgccgtgtacttct 818  
Db 874 aagaagttgcataaaagaagccatcagcgccaacttccatagatattgcacactgt 933  
QY 819 tatgaaatgaacactagagcttcttltgtatgtgaactccctcttattctaatgaga 878  
Db 934 ggtctatgtgtcttaagcactgtgcatctgtgtatgggaactccttgcct-----ctc 987  
QY 879 acctgtgataccatcgaggactgttctgtctgttcttcttaagtgtaaccatagcagtt 938  
Db 988 aatgtaatattctattcttgaaagaagtgttaacgccttcttcttctatttatttgggactt 1047  
QY 939 ttcgatttgagcagcagctccctactttgaacacttcgcaatgcgcagagagctgctt 998  
Db 1048 caglatctgacatttgccccaacaaatagaagccttgcacaaatgcgaagaggcgagacta 1107  
QY 999 tcatatttccaggttatttataagaaccagtatagataacttctccacagcttgata 1058  
Db 1108 tgaatacttcaagaataatttgataatagccaagcattgacagcttcccaacgaaggaca 1167  
QY 1059 taaacctgaatccatagaagaagactgttgaaatttaaaatgttcttccaattatccatc 1118  
Db 1168 caaacccagacagataattggaatttgaattttaaataatgttacttaactaacatccatc 1227  
QY 1119 aagacacatcatcaagaatttctgaaagttcgtaatctcagaattaaagcttggaagagact 1178  
Db 1228 acgaagtgaagttaaagatcttctgaaggcctcaacactgaaggctgaaggcgagagagct 1287  
QY 1179 cgccttggtgtgtctcaatgtgcagtggaagatagctgtacagctgttcgagagct 1238  
Db 1288 agcccttggtgtgcacaacgtgtgctgtggaagaagacacacgttccagctgttcgagagct 1347  
QY 1239 atatgacccgagatgtatgtgtcttcaatgtgtgataagatgaacatcagagctttaaact 1298  
Db 1348 ctacgaccacatagaaggtgcaggtatcagacagacagacagacatccatcaatgt 1407  
QY 1299 ggcgcatatcagagacatttggaggtgtatgcaagagcgtgttcttcttgcgagacac 1358  
Db 1408 gaaggtatctgcgggaaatcatttgggtgtgtgagtcaggaacccgctgtcttgcacacac 1467  
QY 1359 catcaatgaacatatacaagatatacagacagatgtatgcagatbaaagaatgagagac 1418  
Db 1468 gatttgcgaaacacttgcctatgtgcgagaaacgttcaacatgtagagataaggaaagc 1527  
QY 1419 agcaagggaagcaaatgtcgtatgatttcatatgaggtttcccttaataattatcatc 1478  
Db 1528 tgtcaaggaagccaatgtctatgacttcatcatgataaactgcgccacaatttaacacct 1587  
QY 1479 ggtatagggaagaagagatcaaatgagtgtgagggcagaaacagaggttcgcaattgtctg 1538  
Db 1588 ggttgtgtgagagagggcgagcagctgagtggtgggagacagaagaagagatcgcacattgcgcg 1647  
QY 1539 tgccttgatcgaacaccgaagatctgattttagatgaggtcagctgcctgcgagattc 1598  
Db 1648 ggccttgatcgaacaccgaagatccttctgtgtgaggtgcgcagccgctcagacttggagac 1707  
QY 1599 agaaagcaagtcagctgttcaagcttgacgttgagagagcgagcaaggttcgactacaat 1658  
Db 1708 agaaagcaagcgtgtgtgtgcgcgtctgtgataagctagagaagccgagacacat 1767  
QY 1659 cgtgtgtgacacacgacttctactatctgaagtgcagatttgaattgtgacctcaagga 1718  
Db 1768 tgtatagcttcaacgctgttcatagtgccagtgccagctgtcatctgtgtgtttgaggg 1827  
QY 1719 tggagatctgctgaggaagaagagacatagctgaactaaatgccaagaaggttcatata 1778  
Db 1828 tgggtgtcatctgtgaggaagaataatcatgaaagctcatgaaagagagaggttacttact 1887  
QY 1779 ttcactgtgtgttgcacagatatataaaagactgatgaa--cagatggaatcaatgatcat 1837

Db 1888 caaactgtcatagacacagacttagaggaatgaaatgtgacccgagaaataatgtctatga 1947  
QY 1838 attctatgaagaagaagacccaactcacttccctctgtcactctgtgaaagagcatcaagtcag 1897  
Db 1948 atcccaagtgaaactcgtgacctgtgagtgtacttcaagaanaatcaaatcttcttaat 2007  
QY 1898 acttcatatgaagaagcttgaggaatccaccaactta----- 1934  
Db 2008 aagaagatcaatctcgagaaglatccacagagaagacagaccagagagagagacttagtc 2067  
QY 1935 -----agaataagttctctgaagctctctctatataaattttaaagttaa 1982  
Db 2068 gaaagagagtgatgagaagaatgtgctatagtttcccttcttgcacatcctaagctaa 2127  
QY 1983 caagctgaaatgagccttctgtgtctgtgcggaacatgtgctctgtcttaagtgaactgt 2042  
Db 2128 tatatgtgaatgagccatatttgaattgtgtgtgacttgtgcgttataaatgtgtgcat 2187  
QY 2043 tcatcagatatttccatcatccttctgcacaaatataaccaatgtt---ggaataatga 2099  
Db 2188 acacaagttgttgcacatagtttcttcaagaagtgtaggggttcttcaagaagacagca 2247  
QY 2100 taaacacatataagcatatgacatgacagaatttctcatatatttgcataatttgggtgt 2159  
Db 2248 ccatgaaacccaacaaaggaatgtlaactgttctccctcttcttctgtcatgggaat 2307  
QY 2160 tatttgccttgcagatttatttcatatgcagagatttatttcaagcagagcaagggaatatt 2219  
Db 2308 gatttcttctgttcaagacttcttcaagcttctcaagcttcaacatttgcgaagctgagagatcct 2367  
QY 2220 aacgatgagatgaagacacttgcccttcaagcagatgttatcaagatatttccctgtgt 2279  
Db 2368 caccaaagcagctcagatacatgttcttcaaatccatgtctgcagacatgataaagctgtgt 2427  
QY 2280 tgaatgaagaagaacagcagcagagagcttgcacacacatatattgcacatagatatagaca 2339  
Db 2428 tgaagacataaanaaacccctgtgtcgtcgtacacacagagctcgtgtatgacgttcttaa 2487  
QY 2340 aatcgaagagcaaacaggttccagagattgtgtcttcaacacaaatgcaactaaatgag 2399  
Db 2488 tgttaaagggtgtatgtgtcccaagctgtcgtgtgtatccacagatgttagcaaaccttgg 2547  
QY 2400 accttcaatcatcttcccttataatagtatggagagatgaacatctctgttctcgaat 2459  
Db 2548 cacaagatattatcttacttactatgtcttctgtcagcttcaacttacttacttgaat 2607  
QY 2460 tgcctcagactgtcgttgcaggaatgatatgaacccgacagcaatgacgtgatttgcacaa 2519  
Db 2608 tataccactcatgttctgtgtgtgaattatgtgaatgaactgtgtctgttgcgaagcctc 2667  
QY 2520 caaagataagcaagaacttgaatgacatgtctggaagaatgacaaatgaaacttggagaaat 2579  
Db 2668 gaaggaacaagaagaagcttagagatcctcgtggaagaatgcgtacagaagaacattgaaact 2727  
QY 2580 acgtactatagtgatcataacaagaaggaagaaccccttcgagcaaatgtatgaagagatgc 2639  
Db 2728 ccgacagtggttcttcttactgtactagtagagcagaagtttgaactatgtatgtcccgagct 2787  
QY 2640 tcaaacatcaacacgaataactctcgaagaagaagcagatatttgaagcgtgttgcatt 2699  
Db 2788 gcaatatccacacagaagaatgtcttgaagaagaagcacaagcttctgtgatacactcgcct 2847  
QY 2700 cagcagcttcttataatttcttgccttgcagcaggggttctgcatttggagactttaat 2759  
Db 2848 caccagagcagatatttatttcttctatagcgtgtgttccggttgcgttgcacttgtgt 2907  
QY 2760 tcaagcttgagcaatgacccagagggagatgttcatagatttctacgtacatgatatagg 2819  
Db 2908 ggcacaggaactcatgacgttgaatgaatgtatgtgtatctgtctgtgtctgtgt 2967  
QY 2820 agctatggtccttgcgaanaaacgcttcttgcctcctgaatatctccaaagcacaatcggtg 2879

Db 2968 tgcacatgagcaggaataccagttacatctcgtctcgtactacgagcgaagcccaagttctc 3027  
QY 2880 ggtctgcacatctgtttcgtctgttggaaaagaacccaataatagacagccgacgtcaaga 2939  
Db 3028 agcatccacacatcatcagatcatcttgagaataatcccgagatttgacagctacagcaagca 3087  
QY 2940 agggaaaaagccagacacatctggaaggaatttagagtttcgagaagctcttctctta 2999  
Db 3088 gggcttgaagacttaatttgtagaagaaatctgaaatlaattgaatgagatcgttcaacta 3147  
QY 3000 tccatgtgcccagatgtttccatccctcgtggttaccctcagttacgattgagcgagga 3059  
Db 3148 tccacaccgacccaacatcccaagtgcttcaaggaactgagcttgcgaggtgaaagagggca 3207  
QY 3060 gacagtagacatttggggagcagcgctgtgggaaaagcaacttctgttcaacttgcga 3119  
Db 3208 aacgttcgctgtggtgagcagcagtgctgcggagagtagtacagtggtccagctctcga 3267  
QY 3120 gagacttatgaccccgagcaggaacagtgctgtttagtggtagtggtagtgcacaaagaatt 3179  
Db 3268 ggcgtctacacaccccaatggtctggaacagtgcttctagatggcaaaagaataaact 3327  
QY 3180 gaattacagtggtctcgtctcccaatagcaactgtctcccaagaagcctgtgcttcaa 3239  
Db 3328 caacgttcagtggtctcgtcgcgccacttggcatgtgtgtccagagsgccatctgttga 3387  
QY 3240 ctgcagcatctgtgagaacatctgctatgtgtgacacacagcggtgtgtgcaattagatga 3299  
Db 3388 ctgcagcatctcagcagacatctgctacgagacacacgctgtgtgtctcatgagga 3447  
QY 3300 gatcaaaaagccgaaatgtcagcaacatatactcttctttagaagttctccctgagaa 3359  
Db 3448 gatcgtgagggccgcaaggagggccacacatccacagtttcaactacactcgtgagaa 3507  
QY 3360 atacaacacacaaagtctgagcctgaagagcagccttctcgtgcgcgaacaaagact 3419  
Db 3508 atacaacacacaaagtctgagcctgaagagcagccttctcgtgcgcgaacaaagact 3567  
QY 3420 agctattgcaagggctcttctccaaaacccaataatttattgtgtgagtgagcacttc 3479  
Db 3568 cgcacatcgcgccgtccgtctcagacagcctacatcttacttctgtagtgaagcagacatc 3627  
QY 3480 agcctcgtatagatgacatgtggaagtggtttagacatgaccttgatgaagcagaaggg 3539  
Db 3628 agctctgtgacgagagatggaagagtgctcgaagaaagcgtggaacaaagccagagga 3687  
QY 3540 aagagcagcctagtggtcactcagcagcgtctcgaattcagaagcagactttagatgt 3599  
Db 3688 ccgcacctgcatctgtgacgcgcacgcctgttccacatccagaacgcagactttagatgt 3747  
QY 3600 ggttctgacaaatggaagaaatgaagaaacaaagaaatcacaagaagcttctgagaatgc 3659  
Db 3748 ggtgttcaagaagcaggtcaagagagacagcagccacagcagctgtgtgcccagaa 3807  
QY 3660 agacatatatttaagttagtgaatgacagtcagtcagtgta 3702  
Db 3808 aggcatactattctcgtatggttcaaggtgagcgtggaacagcgctca 3850

RESULT 11  
US-08-461-823-1  
Sequence 1, Application US/08461823  
Patent No. 5593840  
GENERAL INFORMATION:  
APPLICANT: Bhatnagar, Satish K.  
APPLICANT: George Jr., Albert L.  
APPLICANT: Nazarenko, Irina  
TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OncorPharm, Inc.  
STREET: 200 Perry Parkway  
CITY: Gaithersburg

STATE: Maryland  
COUNTRY: USA  
ZIP: 20877  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,823  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/168,621  
FILING DATE: 16-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/010,433  
FILING DATE: 27-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Karta, Glenn E.  
REGISTRATION NUMBER: 30,649  
REFERENCE/DOCKET NUMBER: PA-0012 CIP 2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301 527-2058  
TELEFAX: 301 208-6997  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2726 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-461-823-1

Query Match 23.1% Score 856; DB 1: Length 2726;  
Best Local Similarity 61.7% Pred. No. 4.2e-246; Indels 42; Gaps 3;  
Matches 1438; Conservative 0; Mismatches 850;

QY 1407 gatggagagagcagcagaaggaagcaaatgctgatatcttcatcagtggaattcccaataa 1466  
Db 1 GATTGAGAAAGGTGTCAAGAGAGCCATGCTTATCATGTAATGCAATGCCCTCATATA 60  
QY 1467 attatacatgtgtgtaggggaaagagctcaaatgagtgaggagcagaacagagat 1526  
Db 61 ATTGACACCCGTGGTTGAGAGAGAGGGCCAGTTGAGTGGTGAGCAAGCAGAGCAT 120  
QY 1527 cgeaatgtcgtgctccttagtgcgaaccccaagaattctgattttagatgaggtcaagtc 1586  
Db 121 CCCCATTTGACGTGCCCTGGTTGCGAACCCCAAGATCCTCTGTTGATGAGGCGCACGTC 180  
QY 1587 tgccttgatcagaagaagcaagtcagctgttcaagctgcagcggagaaagcgagcaag 1646  
Db 181 AGCTTTGGACACAGAAAGGAGGAGCAGTGTTCAGTGGCTCTGAGTAAGCCAGAAAAG 240  
QY 1647 tggactacaatcgtgtgtgtagcagacacgacttctactatctcgaaagtgagattgtgt 1706  
Db 241 TCGGACACCAATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
QY 1707 gaccctaaagatggaatgtcgtgcgagagaagagacactgtctgaactaatgccaag 1766  
Db 301 TCGTTTGGATGATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
QY 1767 aggtcattatctacactgtgtagtgcacaggaatlaaaagcgtgatgaacagaatga 1826  
Db 361 AGGCATTTACTTCAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 420  
QY 1827 gtcaatgacatactta-----ctgaagaagaagcaactcactct 1869  
Db 421 TGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480





COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,545  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEetical: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4002  
US-08-996-545-1

Query Match 11.5%; Score 426.6; DB 2; Length 4002;  
Best Local Similarity 48.3%; Pred. No. 2.5e-118;  
Matches 1467; Conservative 0; Mismatches 1469; Indels 99; Gaps 6;

Qy 749 aagatgaagagatttggcctaataaaggactatagcttcaaaagtgtcttggctg 808  
Db 956 ACGAAGCTGAGAAATGGGGACAAAGAACCAAGATTGTCATGGGTTTCATGATTGGCGCA 1015  
Qy 809 tgaacttcttaataatgaatgaactatgactgtcttcttggatgtaaacctctgattc 868  
Db 1016 TGTTGGCTTATGACTGCAACTAGCGTCTGGCTTGGATGGGTTCTGCTTCTCTGG 1075  
Qy 869 ttaatggaacactgataataccatcggaactgtcttcttggatgtaaacctctgattc 928  
Db 1076 TAGATGG-----TGCATCTGATGTGGGTGATATCTCACAGTTCTCATGGCATCTTGA 1129  
Qy 929 atagcagttatgcatctggagcagcagctccactcttgaacacttgcgaatagccgag 988  
Db 1130 TCGATCTGCTCTCTGGGGAACGTTAGTCCAAATGCTCAAGCATTTCAAAAGCGCTGG 1189  
Qy 989 gagcgtcttcataatcttcaaggtatgtatgaagaacccagatataaactttcca 1048  
Db 1190 CCGGCGCCGCAAAAGATTTTGGACGATCGATCGCATCCCATTAATCCATTTTGA 1249  
Qy 1049 cagctggaataaactgaatccatagaagaactgtgaaatttaaaatgttcttcca 1108  
Db 1250 ACGAAGGGAAGACGTCGACCATTTTGAAGGCGCAATTGAGTACGCAATGTCAGCAT 1309  
Qy 1109 attatccatcaaacatcatcaaatcttgaaggtctgaacttcaaatgaatgaatctg 1168  
Db 1310 TTTACCCCTCTAGACCCGAGGTACCGTCATGAGAGATGTTCTCTGTCGTAAGCCGCTGG 1369  
Qy 1169 gagagagagctgctctgttgcgttcaatggcagtggaagaagtacggtatgcagcttc 1228  
Db 1370 GAAAAACACCGCTTATGCGCCCTCTGCTGTGAAAAAGTACGGTGGGCTTGG 1429  
Qy 1229 tgcagaggtatagatcgagatgagcttatcatcatgtgtagatagaatgacatcagag 1288  
Db 1430 TTGACGATTTCTACATGCTGTGCGGCTAGCGTTTCTGCTGATGCGCATGACATCAAG 1489  
Qy 1289 cttaaatatgctgagcatatcgagacatattggagtggttaagaaagcctgtttgt 1348  
Db 1490 ACGTCAATCTCGGCTGCTGCGCAACAGATCTTTGTTAGCCAGAGCGCTGTTCTTT 1549

Qy 1349 tcggagaccaccatcaagtaacaatat-----caagatag 1381  
Db 1550 TTGGCAGCAGCATTTATTAAGATATTAGGCACGCTCTATCGGCACAAAGTACGAGAAAG 1609  
Qy 1382 gacgagatgattgactgataagagatggagagacagcaaggagcaaatgtgtag 1441  
Db 1610 AATCGAGGATTAAGGTCCGGGAACATCATCGNAGACGGCGCAAAATGGGATGTGCTCATG 1669  
Qy 1442 attatcaatgagtttccctaataaattatcatattgtaggggaaaaaggagctcaa 1501  
Db 1670 ACTTTATTAAGTCCCTTGGCTGGAAGGTTATGAGACCAATGTTGGGAGCGCTGCTTCTCC 1729  
Qy 1502 tgaagtggaggggcaaaaagagagatgcgaattgcccgtgaccttgaatgcgaaccccaaga 1561  
Db 1730 TTTGAGGTGGCCAGAAACGCGCATTTGCAATGCGCGTGCCTGTGTTATGACCCCAAAA 1789  
Qy 1562 ttctgattatgataagctgaactgtctgccttgccttgccttgccttgccttgccttgcct 1621  
Db 1790 TCTCTCTCTGATGATGAGGTACTTGGCCTTGGACACAAATCCGAGGCGTGTTCAG 1849  
Qy 1622 ctgcactggagaaagcgagcaaaagtcggaactacaatcgtagtagcaacagacttcta 1681  
Db 1850 CAGCTTTGGAGAGGCGCAGCTGGAAGGCCGAACTATTTGTGATCGCTCATCCCTTCCA 1909  
Qy 1682 ctattcgaagtgcaagatttgaattgtgaaccttaaaagatggaatgtctgctggagaaagag 1741  
Db 1910 CGATCAAAACGGCGCAACCATTTGCTTCTGTCATAGGCAAAATGCTGAACAAAGAA 1969  
Qy 1742 caaatgcaacttaattgcaaaaacgagctcatatattcaacttctgtagtcaagagata 1801  
Db 1970 CTCAGATGAATTTGTTGATACCGCGGAGCGGCTTATCGCAAACTTTGGAGGCTCAACGTA 2029  
Qy 1802 ttaaaaa-----gctgatatcaaatgagatg 1828  
Db 2030 TCAATGACAGAAAGAGCTGACGCTTGGAGAGCGCGACCTTATGAGATCTCAAGAAATG 2089  
Qy 1829 caatgacatattctactgaagaagaacacttcttcttgcacttctgtaagagata 1888  
Db 2090 CAGATTTTGGCCAAATTCAAATTCGCTCAAGGCGCATTCGATCGATCGACGGAACCA 2149  
Qy 1889 tcaagtcaacttcatctgacaagagcttag---gaatccacccaactcaaaagatgaagtc 1945  
Db 2150 CAACGATTGACCGGACGCGGACCCACCAAGTGTCTTCCAGCGCGATTTCTTTAAAGAC 2209  
Qy 1946 ttctgtgaagttctctatataaaatttaag-----ttaa 1981  
Db 2210 CCCCCGAAACAACCTCGAAATCTCATATGAGACGCTGCTCAAAATTTGTTGCTTCTCA 2269  
Qy 1982 acaagcctgaatgagccttctgttgcgttgggagacttgccttgccttgccttgccttgc 2041  
Db 2270 ACCGCCCTGAATVCCGTACATGCTCATGCGTCTTGTCTTCTCAGTGTTAGCTGGTG 2329  
Qy 2042 ttcatccagatatttccatcatcttgcgaataataataa-----tgttggaaata 2095  
Db 2330 GCCAACCCACGACAGCAAGTGTATGCTAAAGCCATGACGACACTCTGCTCCAGAAAT 2389  
Qy 2096 atgataaaacacatlaaagcatgataagaaattatccatgatatctgcatttgg 2155  
Db 2390 CACATATATGCAAGCTTGACATGATGCGGATTTCTGCGCATTTGATGTTCTTGGTGG 2449  
Qy 2156 ggttatttgccttctgcttgcgttcttcaatgcagaggaattatlaagcagagagagga 2215  
Db 2450 GTATATTTCAGTTTATTCACGAGTCAACCAATGAGTGTGCTGATTTCCCGATCTCCGAGA 2509  
Qy 2216 tttaacagatgataagtaacacttgccttcaaaagcaatgtatatacagatattgct 2275  
Db 2510 GACTTATTCGTGCGCGAGAACACTGCTTTCGAGACGATCTCCGTCAAGCATTTGCT 2569  
Qy 2276 gtttgaatgaagaagaacagcagcagagagcttgcacaaatattagccatagatatag 2335  
Db 2570 TCTTTGACAGAGAGAAATATGACCGCGGCTGTACGCTTCTTCTGTCACCGAGACGA 2629  
Qy 2336 caaaatccaaggaagcaaacaggttccagagattggtgcttcaacaaataatgcaactaaca 2395



Db	2650	AGCATCTCTCGGGTGTAGGGGTGACCTTAGGCACATCTGTATGACCTCCAGACC	2689
Oy	2396	tgggacttcaatcatcattcccttatatatagatgagagatgacattccctgatctga	2455
Db	2690	TAGGACGGGCTATCATATTATTTGGCCCTGGCATTTGGGGTGGAAATTTGGCTTATTTGTATCT	2749
Oy	2456	gtatgtccagtaacttgcgcgtgacagaaatgatgaaaccgcagcaatgactgatatt	2515
Db	2750	CGGTTGTGCCGGTTCTTCCTG6CATGGGGTTTCTACCGATTCTATTTGCTAGCCAGTTTC	2809
Oy	2516	ccaacaaagataagcaagaacttaagcatbctggaagaatgacaaatgaaacttltgaga	2575
Db	2810	AATCACCTCCAAAGCTTGTCTTATGAGGAGTATTCGAAACTTTCTTCGACGGCTACATAGT	2869
Oy	2576	ataaactactaataagttatatatacaaggaaaaagcccttgagcaaatgtatatgaaga	2655
Db	2870	CTATCCCAAGTTCCTCTTAACTTAACTCCGGGAAAGGATGCTGGGAGATTATACCATGCCC	2929
Oy	2636	tgctcagactcaacaagaataaactctcgaaagaacacagatatltgaagctgtatg	2695
Db	2930	AGCTTTAGCGACAAAGGAGGACCAAGCTATATCTGTCTGTGAGTCATACCTGTTATATG	2989
Oy	2686	catcagccatgccttatatatatttgcctatgacagagglltccatttggagccatt	2755
Db	2990	CGTGTGGCAGGACACTTGTTCCTGTGGTGGCCGCTCGGTTTTGGTACGAGGGACAC	3049
Oy	2756	taattcaagcttgaacgaatgaccccaaggagcaatgtctatgttctcgaattgat	2815
Db	3050	TTCTTGGTCAACAGAGTATGACATTTTCCGCTCTTCTTTGTGTTTTCGCCAGATTCT	3109
Oy	2816	atggagctatgcccatacggaanaaacgcgttltgtcttcgtaatctccaaagccaat	2875
Db	3110	TTGGTGTCTATCCGCGGCGACCGTCTTTTCCTTTGCCACACATGAGCGCAAGCGGAAGA	3169
Oy	2876	cagggggttgcacatcgttltgccttlttgaaaagaaccaataatagaagccgcagtc	2935
Db	3170	ATGCGCGCCCGAATTCCGACACTGTTGACCGGAAGCCACAAATTGATCTACTGCTG	3229
Oy	2936	aagaaggaaanaagccagacacatgtgaagggaaattagaatttcgaagaatccttct	2995
Db	3230	AAGAGGGCGAAGAACTCGAAGACGATGGAAAGTGAATTTATGGAACGTCACCTTCA	3289
Oy	2996	tctatcatgtcgcccaagatgttltatctctcgtgtgtatccctcagtatgagcgag	3055
Db	3290	GATACCCGACCCGCCGAACACGCTGTCTCGCGGCTTGAACCTGACCGGTGAAGCCTG	3349
Oy	3056	gaagaacagtgaatcttgttgggagcagcgcgtgtggaagaacactctgttcaacttc	3115
Db	3350	GACAAATATGTTGGCTTGTGGGACCAAGGGGTGTGGCAAGATACACACATTCGATTGC	3409
Oy	3116	tgcagagacttatgaccccgctgaagaagagctgtcttgaatgtgtgtatgacaaag	3175
Db	3410	TTGAGCCCTTTTACGATGGCATATGGCCGGGTCCATCTCTTTATGGAAGGACATTACTA	3469
Oy	3176	aattgaatgtacagttgctcgttcccaaatagcaatgttctcctcaagaccctgtcct	3235
Db	3470	AACCTAATATCACTCTTACCGGACGCTTCTGTCTACCTGTTCAAGGACCGACACTCT	3529
Oy	3236	tcaacttgaagcatbctgtgaagaacatgcgcbatgtgacacaagccggttgtgtccattag	3295
Db	3530	ACCGAGGCGACCATCAAGGAACAACTCTTACTTGTGTGTCGAAGATGACGATCGGGAAG	3589
Oy	3296	atgagatcaagaagaagccgcaaatgtcagcgaatatctcatcttltatgaaagttccctg	3355
Db	3590	AATCTTGTATTAAAGCTTGCAGAGGACCTTAATTTTCACACTTCATCATGTGCTCCGG	3649
Oy	3356	agaaataacaacacaaagttgagactgaagaagcaacagtttctgtgcgcgcaagaacaa	3415
Db	3650	AGGGCTTTAATACAGTTGTTGGCAGCAAGGAGCAATGTGTTGCTGGCGCCAAAGCAAC	3709
Oy	3416	gaactagctatgcaagggtcttctccaaaaccaaaatttatltgtgtgataggcgca	3475

Db	3710	GTGAGGCATTGGCCAGAGCCCTTCTTGGGGATGCCAAAATCCTTCTTCGATGAAGCA	3769
Qy	3476	cttaagccctcgatatabacagtgaagaagtggtltcagcatgccccttataaagccagga	3555
Db	3770	CGTAGCCCTCGATCTCCGATCTCAAAAAGGTGCTCCAGCGGCTTGGATGCGCGTTGCC	3829
Qy	3536	cggaaagacatgcctcagtggtgtaccccaaggtctctcgaaattgaagcaagcagattga	3595
Db	3830	GAGGCCAGACCAATGCGCGCTTGCCACACCGACTCAGCAAGATTAAAAAGCGGACGTTA	3889
Qy	3596	tagtgtgtctcgacaatgaaagaataaagacaagaaactatcaagagctccctgagaa	3655
Db	3890	TCTATGTTTTCGACCAAGGCAAGATGCTGCGAAAGCGGACGACACAGCACTGTCGAGA	3949
Qy	3656	atcgagacatatatttaagtttgtaatgcaag	3690
Db	3950	AAAAGGCGGTACTACGACCTGTCACACTTGCG	3984

```

RESULT 13
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928698
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atld of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-996-545-3

```

Oy	809	tgcactcttattgaatgtaaacctatgaaactatgaaactgtcttcttttgtaagaaacctcttgattcc	868
Db	1016	UGUUGGCCUUAAUGUACUGAAUADACGGCUTUGGGUUGUGGUGGUCUUGUUGCCUG	1075
Oy	869	ttaatgagaaccttgatataccatcggagctcttcgtctgtttctttagtgtaatcc	928
Db	1076	UAGAUUG-----UGCAGUCCAGUUGUGGUGAAUUCUCACAAUUCUUCUAGGCCAUUCUUGA	1129
Oy	929	atagcagttatctgcatctgagcagcagctccctactcttgaaaccttcgaaatagcccgag	988
Db	1130	UCCGAAUUGUUUCUUGGGGGAACGUAUAGUCUCAAUUGCUCACAGCAUUVUCAAACGUCUGG	1189
Oy	989	gagctgcctcttcattcttccagtgatttgataagaacccagatataagattcttcca	1048
Db	1190	CCGGGGCCCGAAAGAAUUDUGGACAGUADGAGUCCGACGUCUCCCAUUAACUUCUAAUUDUGA	1249
Oy	1049	cagcttgatataaaaccttgataccatataagaagagctgtgaaatttaaaatgttcttcca	11081
Db	1250	ACGAAAGGAAGACGCGUCGACCAUUVUUGAGGGCCACAUUGAGUUAACGAUUAACGA	1309
Oy	1109	attatccataagaacacataltcatcaagatctctgaaaggtcttgaaattcagaattagctg	1168
Db	1310	UUUAAACCAUUAACACCCGAGGUCACCGUADUGAGAGAGUUDUUCUUGCAUAGCCCGUG	1369
Oy	1169	gagagcagctcgagctctgtgctgtcctaactgagcagctggagaagtagatagcttcagcttc	1228
Db	1370	GAAAAACACACCGCUUUAUGUGGCCUUCUGCUCUGGAAAGAAAGUAGGUGGUCGCGUGG	1429
Oy	1229	tgcagagttatatatgtatccgagatgtgcttatacgtgtgatagtgagaatgacatcagag	1288
Db	1430	UUGAGCCAUUDUACAUUGCCUGUCCGGUAGCGUUUUGUCUGAUGCCCAUACUAAAG	1489
Oy	1289	ctttaaatgtgaggagcatlaccgagaccatattgagtgtagtaagaagctgttctgt	1348
Db	1490	ACCUCAAUUCCCGUGCGUUGCGCCACACAAUUCUUUGUGUAGCCAGAGCCUUGUUCUUU	1549
Oy	1349	tcggagaccacatcagttaacaatat-----caagtat	1381
Db	1550	UUGGCACGACGAAUUUUAAGAAUAUUAAGCAGCGUUCUACUGGCACAAAGUACGAAUG	1609
Oy	1382	gagcagatgtatgtgactatgaaagagatgagagagcagaagaggaagcaaatgcgtatg	1441
Db	1610	AUUDCCGAGGAUAAAGUUCGCGGAACUCUADUGCAACGCGGCAAAAAGCGGAUUGCUUGC	1669
Oy	1442	attttatcatggaaggtttctccataaatttaatacattgtgtaagggaataaagagctcaca	1501
Db	1670	ACUUUAUUAUCUGCCUUGCCUGAAGUUUAUAGAACCAUUSUUGGACGUGGUCUUUCCU	1729
Oy	1502	tgaagtgaagagcagaaacacagagatctgcaaatgtgctgttcttgatctcgaaaccccaaga	1561
Db	1730	UUUDCAGGUGGCGCCAGAAACACGCAUUGCAAUUGGCGCCGUGCCUUUUGAGUACCCAAAAA	1789
Oy	1562	ttctgatttaagttagtgagctcagctgtgcgcctgattccagaagcaagctcagctgttcacg	1621
Db	1790	UCCUGCUCUCCGGAAGAGCUAAUUCUGGCCUUGGACACAAAUUCCGAAGCGUGGUCUACAG	1849
Oy	1622	ctgacactggaagagcgagagcaaaaggtctcggaactacaactcgtgtgtgaacacacgacttcta	1681
Db	1850	CAGCUUUGAGAGGCGCAGCUGAAGGCGCGAAUUAUUVUUGAUUGCUGUACUUGCCUUGCA	1909
Oy	1682	ctatctgaagtgtcgaatttgatgtgtgacaccttaagatgtgaatgtctctgctcgaggaagagag	1741
Db	1910	CGAUCAAACCGCGCCACAAACAUUUGUGUUCUGUCAAUUGGCAAAAUUGCUGAACAAAGAA	1969
Oy	1742	cacatgtcgaactaatgtgcaaaaagaggtctatatattcatcttctgtatctgatacagagta	1801
Db	1970	CUCACGAAUUAUUGUGUAGACCGGCGAGGCGCUUUAUGCAAAUUVUUGAGGUGCUCAAAGUA	2029
Oy	1802	ttaaaaa-----gctatgaaacagatgtag	1828
Db	2030	UCAUUGAAGAAAGAGAGCUGACCGCCUUGAGAGCGCCGACGCGCUGGAGUUCUACAGAAUG	2089

QY	1829	caatgcacatattctacttgaaagaagaaccaactcaattccctctgacatctgtgaaagga	1888
Db	2090	CAGAAUUAUUGCCAAAUUCCAAAUACGCGUCAAAGCGCAUCAAUCCGAUUCACACGGAATAACCA	2149
QY	1889	tcaagtcagaacttcatttgaaagaagctgag---gaatccaccacactaaagagataagtc	1945
Db	2150	CAACCAUUGACCCGACGGGACACCCACAAGUUCUGUUUCCAGCGCGAUAUUCUUAAGAAAGAC	2209
QY	1946	ttcctgaagctcctctatttaaaattttaag-----ttaa	1981
Db	2210	CCCCGAAACAAACUCCGAAAUUACUUAUUVGACGCGUCCUCAAUUAUUUUGUCCUUCUUA	2269
QY	1982	acaaagcctgaatgagccttttgtgtgttcctgggaacatitgctctgttcttaatggaactg	2041
Db	2270	ACCCCGCUAAAUUCCCGUACAUUCCUAUUGCGCUUUGUUCUCCAGAGUUAACUGUGUG	2329
QY	2042	ttcctcagatatttccatcatcttgcgtgaaataataacca-----tgtttggaata	2095
Db	2330	GCCAACCCACGCAAGCAGUCCUUAUUGCUAAAGCCAUACAGCACACUUCUCCGCCAGAU	2389
QY	2096	atgataaacacacatlaaagcatgatgacagaataattatctcatctgatatctgtcatttgg	2155
Db	2390	CACAAUUAUAGAACUUCUACAUAGUUGCGGAUUAUUCUGUCAAUUAUUAUUAUUCUUGUGU	2449
QY	2156	gtgtattctgtctgttcagttatctcatgacaggaattatttaagcgagagcaggaggaa	2215
Db	2450	GUAAUUAUUCAGUUUAUUAUCCACAGUCAAACCAUAUUGUUCUUAUUAUUGCCUUAUUGCCGGA	2509
QY	2216	ttttaagatagagttaagacactgttcgttcaagacgatgtatataagatatgtcgt	2275
Db	2510	GACUUAUUCGUCGCGCGAAGACACUUGCCUUCGAGCAUUAUUCUCCGCUAAAGCAUUGCUU	2589
QY	2276	ggtttctgtaaaagaaacagacagagagugctctgacacaaatataagcataagatatag	2335
Db	2570	UCUUUAGCAAGAGAAAGAAUAGCAGCGCGGCUUCUGACUCCUUAUCCUUCACACCGAAGAC	2629
QY	2336	cacaaatcaaggagcaacaggttccaagatttggcglttcaacacaaatgcaactaa	2385
Db	2630	AGCAUUAUUCUCCGCGUUAAGCGGUGUACUUAAGACAGAUCCUUAUAGACUCCACAGACC	2689
QY	2336	ttgggaacttcaagtattcattcccttattatagatgagaggagatgcatctccgattcga	2455
Db	2690	UAGAGCGCGCAUUAUUAUUGCCUUGCGAUUGGUGGGAUUAUUGCCUUAUUGUUAUUCU	2749
QY	2456	gtatgtctccagactgtccgtgacaggaatgataagaaacgacgaatgacatgattgt	2515
Db	2730	CGGUUGUCCCGGUUCCUCCUGGCANUGCGGUGUUAUACCGAUUAUUAUUGACGCCAGUUC	2809
QY	2516	ccaacaagaatagcaagaacttaagcatgtctgaaagaatagcaactgaagctitgaga	2575
Db	2810	AUUCACGCUCCAAACGUUUCUUAUAGAGGAUUCGAAAUUUUUGUCCUUGCGUACUAUCGU	2869
QY	2576	atatacgtactatagtttcatttaacaaaggaaagaccttgcagcaaatgtatgtaagaga	2635
Db	2870	CUAUUCGCGACAGUUGGCUUAUUAUCCGGGAAAGGGAUUGUCUGGGGAUUAUACUAGGCC	2929
QY	2636	tgcttcagagatccaacaagaataacctcgaagaagaacacagatatitggaagctgtat	2695
Db	2930	AGCUUAGCGCAACAGGACGAGCCAGUUAUUAUUCUUGUAGGUCAUCCCGUUAUUAUG	2989
QY	2696	catlcagacagaccttataattttgtcctatgcaagaggttctgatttgaagcctat	2755
Db	2990	CGUGUCCGACGACAUUUGUUUUCUUGCGUGGUGCGUCCGGUUUUGUAGCGAGGAGCAC	3049
QY	2736	taattcaagctgtgacgaatgaaccccaaggagcatgttcatagatttttactgtcaattgat	2815
Db	3050	UUUCUUGGUCACACAGAUUAUGACAUUUUUCGCGUCCUUAUUGUUUUGUCCGGAUUAUCU	3109
QY	2816	atgagagctatggcatcgagaanaacgctgttttggctcctgaaatatttcaagaacaaat	2875
Db	3110	UUUGUGUCCAAUUAUCCGGCGGACCGUUCUUUCCUUAUUGCACCAACAAUUGGCAAGGCGGA	3169
QY	2876	cggaggctgtgcgcatctgttttgccttgtttgaaagaacaaataatagacagcgcgactc	2935

```

Db 3170 AUGCGCGCGCGGAUUCGCGAGCUGUUCGACCGAAAGCCACAAUUGAUACUGUCUG 3229
Oy 2936 aagaagggaaaaagcagacacatgtgaaggatttagatttgagaagctcttct 2995
Db 3230 AAGAGGGCGAGAGCUCGAAACGGUGAGGUAUUGCAUUGAAGACGUCACUUA 3289
Oy 2996 tctaccatgtgtccagatgttttccatccctcgtgtgttaccctcagatgtgagcg 3055
Db 3290 GANACCGGACCGCGGCAAGACAGCCUUCUUCGCGGCUUGAGCUCGACCGGAAAGCCG 3349
Oy 3056 gaaagacagtagcattgttgggagcagcgctgtggaagaagcatttgttaactc 3115
Db 3350 GACAUAUUGUUGCGCUGUGCGACCGCUGUGGCAAGAGUACACCAUUGCAUUGC 3409
Oy 3116 tgcagaagcattatgaccccggtcgaagcaagtgctgtttatgtgtgtgagcaaa 3175
Db 3410 UUGAGCGCUUUGAUGGAGUUGCGGGGUCCAUCCUUGUUGAUGGGAAGACAAAGUA 3469
Oy 3176 aattgaatgacgtggtcgttcccaatagcaatgcttcccaagaagcgtgtct 3235
Db 3470 AACUAAUUAUACUCCUACCCGACUUCUUCACUGUUCAGACGAGCGGACACUUG 3529
Oy 3236 tcaactgcagcattgtctggaacatcgctatgtgtgacacagccgtgtgtgacatag 3295
Db 3530 ACCAGGGCGACCAUACAAGAAACUUAUUGUUGUUGCAAGAUAGCAGACCGAAG 3589
Oy 3296 atgagatcaaaagacccgaatgcagcaaatatccattctttattgaagcttccctg 3355
Db 3590 AAUUCUUGAUUAGGCUUUGCAAGAGCGUAAUUAUUAUUAUUAUUGCUCGCGCG 3649
Oy 3356 aagaatcacacacaaagtgtgacgaagcagcaagcttctctgcgcgcgaagaaca 3415
Db 3650 AGGCGCUUAUUAUACUUGUUGGACGACAAAGGAGCAUUGUUGCUGCGGCCCAAAAGCAAC 3709
Oy 3416 gactgtctatgtcgaaggctcttctccaaaaacccaatttatgtgtgagtgagcca 3475
Db 3710 GUGUGGCGCCUUDGCCGAGCCUUCUUGGAGUCCCAAAUUCUUCUUGCAUAGAGCA 3769
Oy 3476 ctctgcgcctgatatgaacagtgaggaagtggttcacagatgccttctgtaagccaaga 3535
Db 3770 CCUCCAGCCUUCACUCCGAGUACGAAAGGUCGUCGAGCGGCUUUGUUGCGCGGCC 3829
Oy 3536 cggagagagcagtcagtcagtcactcagcagctctctgcaattcgaagcagatttga 3595
Db 3830 GAGGCCGGAACCAUUGCCGUGGACACCGACUCAGCAUUAUAAAGCGGAGCUUA 3889
Oy 3596 tagtgtctgcacaaatggaagataaagaaacaagaaatcatcaagagctcctgagaa 3655
Db 3890 UCUAUGUUUUGACCAAGGCAAGUUGGAAAGCGAAGCGCAAGCAUUGUUCAGA 3949
Oy 3656 atcgaacatatattttaagttgaatgacag 3690
Db 3950 AAAAGGCGGUAUACGAGCUGGUCACUUCGAC 3984

```

```

RESULT 14
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Maard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana

```

```

; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
; US-09-328-320-1

```

```

Query Match 11.5%; Score 426.6; DB 4; Length 4002;
Best Local Similarity 48.3%; Pred. No. 2,5e-118;
Matches 1467; Conservativity 0; Mismatches 1469; Indels 99; Gaps 6;

Oy 749 aagatgcaaaagatttggcataaaagaactatagcttcaaaagtgtcctgtgtcgt 808
Db 956 ACGAAGCTGAGAAATGGGGAACAAAGAACAGATTGTCATGGTTTCATGATTGGCGCA 1015
Oy 809 tttactctttatgaatgaacactatgactgtcttcttgatgagacccctctgattc 868
Db 1016 TGTTCGCTTATGTACTGAACTAGCGTCTTGCTTGAGTGGGTTCTCGTTCTGG 1075
Oy 869 ttaatggagaacctgatataccacoggaactgttctgtcttcttctttagtgaatcc 928
Db 1076 TAGATGG-----TGCAGTCGATGTGGGTGATATTTCTCAGTTCTCATGGCCATCTTGA 1129
Oy 929 atagcagttatgtcattgagcagcagtcctcactttgaaaccttcgcaatagccgag 988
Db 1130 TCGGATCGTTCTCTTGGGGAACGTTAGTCCAAATGCTCAACATTTCAACGCTGTGG 1189
Oy 989 gagcgtccttcatatttccaggttatgtataaagaaccocgatatagaactttcca 1048
Db 1190 CCGCGCGCGCAAAAGATTGGAACGATCGATCGCCATGCCATTATTTGCA 1249
Oy 1049 cagcgtgatataaacctgatatccatagaagaagcgttgaatttaaaatgttcttcca 1108
Db 1250 ACGAAGGGAAGACGCTCGACCAATTTGAGAGCGCCACATTCAGTTACGACATTCAGCA 1309
Oy 1109 attatccatcaagaccatcatalcaagatctgaaagtlcgaatccagaatlaagtcgt 1168
Db 1310 TTACCCATCTAGACCGGAGTCACCGTCATGAGAGATTTCTGTGCAATGCCCGCTG 1369
Oy 1169 gagagacagtcgtcgttgcgttccatgagcagtgagggaagtgatgagtgcaagcttc 1228
Db 1370 GAAAAACAACGCTTAGTGGGCCCTCTGCGCTGTGAAATAAGTACGGGTGCGGCTTGG 1429
Oy 1229 tgcagaaggtatatgtatccgagtgatgtgttatacgtgtgataagtaagcatcagag 1288

```

```

1430 TTGACGATTTCTACATGCTTTCCGGTACGGTTTGGTGGATGGCCATGACATCAAG 1489
QY 1289 ctttaaatgtgycattatcgagaccatatlgagtggttaagtaagaagcctgtttgt 1348
Db 1490 ACCTTAAATCTCGGTGGCTTCCGCAAGATCTTTGGTTAGCCAGGAGCTGTCTTT 1549
QY 1349 tggagcaaccactcaatlaacat-----caagtatg 1381
Db 1550 TTGGACACGAGTTTATTAAGATATTAGGACGGTCTCATCGGCACAAAGTACGAGATG 1609
QY 1382 gacgagatgatlgaactgatgaagatlgagagagcagcaaggaggaatgagatg 1441
Db 1610 AATCGAGAGATTAAGTCCGGGAACATCAGAGACGCGCAAAATATGGCAATGCTCATG 1669
QY 1442 attatacatgagtttccctaaataattatcatattgtaggggaaaaagagctcaaa 1501
Db 1670 ACTTATTAATGCTTGGCTGTAAGGTTATGAGACCAATGTTGGGACGCTGGCTTCTCC 1729
QY 1502 tgagtgagggagagaagaagagagatgcgaattgctgcgttccttagttgaaacccaaga 1561
Db 1730 TTTTCAGGTGGCCAGAAACGCGCATTTGCAATGCGCGTGCCTGTTATGTAACCCAAAA 1789
QY 1562 ttctgatttagatgagagctacgtctgccttgatlcagaagaagcaagtcaagtltcaag 1621
Db 1790 TCTCTCTCTGGATGAGGTACTTTCGGCTTGACACAAAATCCGAGCGTGGTTCAAG 1849
QY 1622 ctgcactgagaagcgagcaaaagtcgagactacaactcgttgtagacaacgaacttcta 1681
Db 1850 CAGCTTTTGGAGGGCAGCTGTAAGGCCGCAATCTATTTGATTCGCTCATCCCTTCCA 1909
QY 1682 ctaattcgaagtcagattgattgtaacccctaaagatgaaatgctgaggaggaagagg 1741
Db 1910 CGATCAAAACGGCGCAACAATTTGGTTCTGTCAATGCAAAATGCTGAACAAAGGA 1969
QY 1742 caacatcgaactaaatgcaaaacgagtlctatattatcacttgatgltcaagata 1801
Db 1970 CTCAGAGATGAATTTGGTTGACCGCGGAGGCGCTTATCGCAAACTTGTGGAGGCTCAACGTA 2029
QY 1802 ttaaaaa-----gctgataacagatgagat 1828
Db 2030 TCATAAGACAGAAAGAGCTGACGCTTGGAGAGCGCCAGCGCTGAGATCTCACGAATG 2089
QY 1829 caatgacatatctctgaagaagaagaccactacttctctgcaacttgtagaagata 1888
Db 2090 CAGATATTGGCCAAATTCAAATCTGCGTCAAGGCAATCAGATCCATTCGACGAGAAACCA 2149
QY 1889 tcaagtcagactlctgaagaagctgag---gaatccaaccaatcctaagaagataagtc 1945
Db 2150 CAACCATTTGACCGCAGCGGACCCCAAGTCTGTTCCAGCGCGATTTCTTTAAAGAGAC 2209
QY 1946 tctcgaagtcctctatataaatttaag-----ttaa 1981
Db 2210 CCCCCGAAACAACTCGAAATCTCATATTGACGCTGCTCAAAATTTGTTGCTTCCCTCA 2269
QY 1982 acaagcctgaatagagcctttgtggttcgaggacattgctctgtctcctaaltgaacg 2041
Db 2270 ACCGCCCTGAAATCCGATACATGCTCANTCGTCTGCTTCTCAAGTTTAACTGAGTGAGTG 2329
QY 2042 ttcaaccgaattlctcaactccttgcaaaaatilaacca-----tgattggaata 2095
Db 2330 GCCAACCCACGAGACAGTGTATATGCTMAAGCCATCGACACATCTGCTCCAGAT 2389
QY 2096 atgataaaaccaatlaagaagcatgtagaataattatcatgatataatcgtcaatttgg 2155
Db 2390 CACAATATATCAAGCTTCACATGATGCGGATTTCTGGCATTTGATGATCTTCTCGGTGG 2449
QY 2156 ggttatttgccttgcaattlctcaatgcaaggaattatttaagcagaagaggga 2215
Db 2450 GATATATTCAGTTTATTCACGCGATCAACCAATGCTGCTCATTTGCGTATGCTCCGAGA 2509
QY 2216 tttaacatgagatlaagacacttgccttcaagccaatgtlatatcaggaattgct 2275
2510 GACTTATTCGTGCGCGAGAGACACTGCTTTGAGAGATCTCCGTCAAGACATTGCTT 2569
QY 2276 ggttgatgaaagaagaacagcaagagagcttgacaacaatatatgacatatagatag 2335
Db 2570 TCTTTGACAGGAAGGAATATACCGCGGCTGTGACCTTTCTCCGTGTCACCGAAGCA 2629
QY 2336 caaaatcaaggaagaagaagtlcccaagatttggcgtlcttaacacaaaatgcaactaca 2395
Db 2630 AGCATCTCTCCGCTGTTACCGGTGACTCTAGGACACGATCTTGATAGACTCCACGACCC 2689
QY 2396 tggagcttcaatlaatcattcccttataatgagatggagatgacatccgtatctga 2455
Db 2690 TAGAGCGGCTATCATTTATGCTTGCCTGGCATTTGGGAATATGGCTTAGTTGATCT 2749
QY 2456 gtaattcctcagacttgcctgcgtgacaaggaatgaattgaacccgaagcaatgacttgg 2515
Db 2750 CGGTGTCGGGTTCCTCGGCAATGCGGTTTACCGAATCTATATGATGACCCACTTTC 2809
QY 2516 ccaacaagaataagcaagaacttaagcatgctlgaaagatgacaactgaagcttggaga 2575
Db 2810 AATCAGCTTCCAAAGCTTGCTTATGAGGATCTGCAAACTTTGCTTCGAGCTACATCGT 2869
QY 2576 atatacgtactatagtlcatlaacaaggaanaagccttcgaagcaaatglatgaagaga 2635
Db 2870 CTATCCGACAGTTGCGTATTAAACCGGGAAGGATGCTGGAGATTTACCATGCCC 2929
QY 2636 tgcctcagaactacaacagaataactcgtgaagaagacagatlatatgaagctgtatg 2695
Db 2930 ACTTGTACCAACAAGGACAGACAGCTATATCTCTCTTGTGAGGTCATCCCTGTATATG 2989
QY 2696 catcaagcatgcccctatataatttgcctatgacagagatttgcatttgaagcctatt 2755
Db 2990 CGTCTGCGAGGACACTTGTCTTCTGCGTTGCGTCCGCGTTTGGTACGAGGACAC 3049
QY 2756 taattcaagctlgacgaatgaccccaagagagcatgltcatagtlttactgcaattgcat 2815
Db 3050 TTTCTTGTGACACGAGATATGACATTTTCCGCTTCTTGTGTTGTTCTCCGAGATCTCTG 3109
QY 2816 atgagactatgcatcggaaaaaagcctcgttttgcctcgtatattccaagaaccaat 2875
Db 3110 TTGTGCTCAATCCGCGGACACCGTCTTCTTTCACACAGACANAGGCAAGGCAAGAA 3169
QY 2876 cgggggcttgcacatcgttctgcttggtaaaagaagaacaatatagacaagcagtc 2935
Db 3170 ATGCGCGCCGAGATTCGACAGCACTGTGACCGAAGGCCAACAAATTTGATTTGCTG 3229
QY 2936 aagaaggaanaaagccaagacacatgtaagaaggaaatttagaagtttcgagaagctcttct 2995
Db 3230 AAGAGGCGAGAAAGCTCGAAAGCGGTGGAAGTGAATTCGAATTTAGGAAGTGCTACTTCA 3289
QY 2996 tctatccatgtgcagatglttcatccctcgttgctatccctcagtatgagcgag 3055
Db 3290 GATATCCGACCGCCGACGAAACGCTGCTGCGCGCTTGGACCTGACCGGTGAACCTG 3349
QY 3056 gaagaagatgaatltgtaggagagcagcgtctgtaggaagaacatctcgttcaacttc 3115
Db 3350 GACAAATATTTGCGCTTGTGCGACCCAGCGGTTGTGCAAGATACACCATTTGCAATTGCG 3409
QY 3116 tgcagaagcttlatgaccccgltgcaaggaacaagtgctgttgaatgltgtagatgcaaaag 3175
Db 3410 TTGAGCGCTTTATGATGAGATTTGCGGCTGATCTCTTGTGATGGGAAGGACATTAAGTA 3469
QY 3176 aattgaatgacagtggtcccgltcccaatagcaatcgltctcaagagcgtgtcct 3235
Db 3470 AACTAATATATCAACTCTACCGAGCTTTCCTGCACTGTGCAACCGAGGACCAACACTGT 3529
QY 3236 tcaactgcagcatgtctgacatcgctatgtagaacaacgctgtgtgtccatag 3295
Db 3530 ACCAGGCGACATCAAGGAAGAAACATTTACTTGTATTTGTCGAAGATGACGTCCGGAAG 3589
QY 3296 atgagatcaaaagaagccgaanaatgacgaataatcatcttatttgaagtlctccgt 3355
Db 3590 AATTTTGAATTAAGCTTCAAGGACGTAATATCTACGACTTCATCATGTCGCTCCCGG 3649
```



Db 1910 CGAUCAAAACGGCGCACACAUUUGUGUUCUAUGGCAAAAUUUGUGCAAGCAAGAA 1969  
Qy 1742 cacatgctgaactaataatgcaaaaacgagtcctatattcactgtgtatgtaacagata 1801  
Db 1970 CUCACAGUAUAUUGUGUACCCGCGAGCGCUUUAUUGCAAAUUGUGGAGCGUACAACGUA 2029  
Qy 1802 ttaaaaaa-----gctgtagaagatgagat 1828  
Db 2030 UCAGUAGAACAGAAAGAGACUGACGCCUUGAGAGACGCCACGCGUGAGUUCACAGAAUG 2089  
Qy 1829 caatgacatattctactgaagaagaagaacactctctctctgcaactgtgtgaagaaga 1888  
Db 2090 CAGAUUUGUCCAAAUAUCAAACUUGUGUCAAAGCGAUCAUCCGUAUCCGACGGAACCA 2149  
Qy 1889 tcaatgactcttctatgcaagagctgag----gaatccaccaactctaaagagataatgc 1945  
Db 2150 CAACCAUUDAGCCGACAGGCGACCCACACAUUCUGUUCAGCGCGAUUCUUUCAAAGAG 2209  
Qy 1946 ttcctgaagtcctctctatataaatttaag-----ttaa 1981  
Db 2210 CCCCCGAAACAAUCUCCGAAAUACUUAUUGGACGCGUCCUAUUUUGUGUCCUUA 2269  
Qy 1982 acaagcctgaatgagccttctgtgtctgtgagacattggtctctgtcttaaatgagatg 2041  
Db 2270 ACCGCGCCUGAAAUUCCCGUAUUGGUCUACUGGUCUUAUUGGUAAGUUGGUGUG 2329  
Qy 2042 ttcacacagatttccatcatccttctgcaaaaataacca-----tgtttggaata 2095  
Db 2330 GCCAACCCACGACAGAGUGCUUAUAGCUMAAACCAUACACACUUCGUCUCCGAAU 2389  
Qy 2096 atgataaaccacatgaagaagcagatgaagaattatctcatgactatgctgacttgc 2155  
Db 2390 CACAAUAUAGCAAGCUUCCAGUAUGGCGAUUUCUGGUAUUGAUGUUCUUGUGUGUG 2449  
Qy 2156 gctgatttgccttctgacttatttcaatgaggaattatttcaagcagagcaggggaa 2215  
Db 2450 GUUUAUACUUAUUGACGACGUAUACCAUUGGUGUUAUUGGCUUAUUGUCCGAGA 2509  
Qy 2216 ttttaacgatgagatgaagaacactgtgcttcaagccatgtatatacagatattgcct 2275  
Db 2510 GACUUAUUGUGCGCGGACAGACACUCCUUDGAGAGUACUCCUCCAAAGAUUGCU 2569  
Qy 2276 ggtttgtaagaagaagaacagcagagagccttgaacaacatattatgcatgatatag 2335  
Db 2570 UCUUUAUACAGAAAGAAUAGACACCGCGCUCUUGACUUCUUCGUCACCGAGACGA 2629  
Qy 2336 cacaacatcaagagcagacaggttccagatgtgcgttcttaacacaaatgtaactaaca 2395  
Db 2630 AGCAUUCUCCGUGUUGAGCGGUGAGUACUAGCAGAUUCUUGACCUCCACGACCC 2689  
Qy 2396 tgggacttcaagtatcatcttcttatataatgtagtggagatgacatctctgattctga 2455  
Db 2690 UAGAGCGGCUUAUUAUUGGCGGCAUUGGUGGAAUUGGCGCUUAUUGUUAUUC 2749  
Qy 2456 gtaatgctcagactgtgcgtgacaggaatgtatgaacccgacgaatgactgagttct 2515  
Db 2750 CGUUAUUGCGGUGUUCUUGGCGACUGCGGUUUAUCCGUAUUAUUGUCCAGUUC 2809  
Qy 2516 ccaacaaataaagaagaacttaagcatgtctggaagaatagcaactgaaacttggaga 2575  
Db 2810 AAUACAGCUCCAAAGCUUUAUUGGAGGAGUACUUAUUGGUGAGGCGUACUUCU 2669  
Qy 2576 atataagtaataatgtatcaataaagaagaagaagccttgcagcaaatgtatgaaaga 2635  
Db 2870 CUAUCCGACAGUUGGCUUAUUAUCCGGAAGGAGUUGGUGGAGUUAUUAUUAUCC 2929  
Qy 2636 tgcctcagactcaacaagaataactcgaagaagaacagactatttgaagaagtgtat 2695  
Db 2930 AGCUUAACGACAGAGGACGACGUAUUCUUCUUGAGGUAUUAUCCUUAUUAUUG 2989  
Qy 2696 caticagcagcttctatataatttgcctatgcaagcaggttctgattgagactatt 2755  
Db 2990 CGUGCGGACGACUUAUUGUUAUUGGUGGCGUUGGAGUUAUUGGAGGAGAC 3049

Qy 2756 taattcaactgagcaatgaccccaagggcatgttcaatgttactgcaattgcat 2815  
Db 3050 UUCUUGGUCACACGAGUAUUAUUCGCUUUAUUGUUAUUCUCCGAGAUUCUUC 3109  
Qy 2816 atggaactatgagcactggaagaagcgtctgttgcctctgataatccaaagccaat 2875  
Db 3110 UUGUGUUAUUAUUGGCGGACCGUUAUUGUUAUUGGUGUUAUUGGUGUUAUUG 3169  
Qy 2876 cggggctgcgaactgttgcctgtgtgaaagaaccacaaatagacagcgcagtc 2935  
Db 3170 AUGCGCGCCCGAAUUAUCCAGACUUAUUGGAGGAGGUAUUAUUGGAGUUAUUG 3229  
Qy 2936 aagaagagaaagaacagacacatgtgaagggaatttagatgttgcagagctctct 2995  
Db 3230 AAGAGGCGAAGACUCCGAAAGUGGAGGUAUUAUUGGAGGAGGUAUUAUUGGAG 3289  
Qy 2996 tctatcactgctgcacagatgttctatcctcgtgtgcttactcactgaattgagcag 3055  
Db 3290 GAUACCGCAACCGCCGAGAAACACUUAUUGGCGGUGUUAUUGGAGUUAUUGGAG 3349  
Qy 3056 gaaagacagtaagcattgtgtggaagcagcgtgtgtggaagaacacttcttcaatc 3115  
Db 3350 GACAAUUAUUGGCGUUAUUGGAGGAGGAGGUAUUAUUGGAGGAGGUAUUAUUGG 3409  
Qy 3116 tgcagagactttagaccccgtagaagcagtgctgtgtgtgtgtgtgtgtgtgt 3175  
Db 3410 UUAAGCGCUUUAUUGGAGGAGUUAUUGGCGGUAUUAUUGGAGGAGGUAUUAUUG 3469  
Qy 3176 aattgaatgtagatgt 3235  
Db 3470 AAUUAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3529  
Qy 3236 tcaactgacagcattgtgtgaacatgcgctatgtgtgaacagcgtgtgtgtgtgt 3295  
Db 3530 ACCAGGCGACCAUAGGAAACAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3589  
Qy 3296 atgagatcaagaagcgcgaatgacgaagaatcttcttatttgaagtgtgtgt 3355  
Db 3590 AAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3649  
Qy 3356 agaatacaacaacacagatgtgaaggaagcagccttctgtgtgtgtgtgtgtgtgt 3415  
Db 3650 AGGCGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3709  
Qy 3416 gactagcatttgaaggtcttctccaaaaccacaaatlttatgtgtgtgtgtgtgt 3475  
Db 3710 GUGUGGCGAUAUUGGCGGAGCCUUAUUGGAGUUAUUAUUAUUAUUAUUAUUA 3769  
Qy 3476 ctccaagcctcgaatgaagcagtgagaggtgtgtgtgtgtgtgtgtgtgtgtgt 3535  
Db 3770 CGUACACCCUCCGACUCCGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3829  
Qy 3536 cgggaagacatgctagtgctcactcaacagcgtctctgtgtgtgtgtgtgtgtgt 3595  
Db 3830 GAGCGGAAACAAUAUUGCGUUGGACACCGAGUUAUUAUUAUUAUUAUUAUUA 3889  
Qy 3596 tagtgttctgcaaatgaaagataaagaagaagaacttcaagaagctctgtgtgt 3655  
Db 3890 UCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3949  
Qy 3656 atcgaacatataatttgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3690  
Db 3950 AAAAGGCGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3984

Search completed: April 22, 2002, 18:43:41  
Job time: 12804 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:32:03 ; Search time 9392.76 Seconds  
(without alignments)  
4235.270 Million cell updates/sec

Title: US-09-873-409-13

Perfect score: 3702

Sequence: 1 ttccgcttgcgtatgcagact.....atgcacagtcagtcagta 3702

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpl:\*  
6: em\_estlba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rnd:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	545	14.7	545	10	AL040762
2	386.6	10.4	943	10	AL520322
3	385.8	10.4	405	10	AA243820
4	378.8	10.2	872	12	AK020318
5	357.4	9.7	559	11	BF692596
6	298.8	8.1	894	11	BF584668
7	275.2	7.4	726	11	BC293345
8	273.2	7.4	1019	11	BG248052
9	269	7.3	944	11	BF796582
10	262.2	7.1	785	10	AV709991
11	253.2	6.8	780	11	BG587938
12	252.4	6.8	560	10	BE749379

13	252	6.8	886	13	BH139685
14	249.4	6.7	981	11	BF969667
15	249.2	6.7	932	13	AZ670821
16	249	6.7	795	11	BF313560
17	248.4	6.7	880	13	AZ687805
18	241.4	6.5	886	13	AZ540627
19	239.6	6.5	871	13	AZ682350
20	239.6	6.5	897	13	AZ541090
21	238.8	6.5	853	13	AZ679807
22	237.2	6.4	834	13	AZ548312
23	234	6.3	947	13	AZ683753
24	232	6.3	503	13	AQ380214
25	230.2	6.2	823	13	AZ532602
26	229.4	6.2	482	11	BI344244
27	227.8	6.2	651	10	BE016688
28	226	6.1	687	11	BG646725
29	225	6.1	919	13	AZ690701
30	224.6	6.1	891	13	AZ682250
31	224.2	6.1	842	13	AZ671925
32	219.8	5.9	661	11	BG302675
33	215.2	5.8	899	13	AZ548799
34	214.4	5.8	921	13	AZ687628
35	213.8	5.8	575	10	AA434644
36	213	5.8	519	10	AA434959
37	212.2	5.7	605	10	AV616675
38	210.4	5.7	750	11	BG585786
39	208	5.6	851	11	BF969062
40	206.4	5.6	803	11	BG298756
41	204.4	5.5	787	11	BG584063
42	199	5.4	854	13	AZ682568
43	197.8	5.3	402	11	H82550
44	197	5.3	541	10	AI722858
45	196.4	5.3	2676	12	AK014319

#### ALIGNMENTS

RESULT 1  
LOCUS AL040762 545 bp mRNA  
DEFINITION DKFZp434C1815.r1.434 (synonym: htes3) Homo sapiens cDNA clone  
ACCESSION AL040762.1 GI:5409708  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 545)  
Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
Wiemann, S.  
EST (Bloecker, et al.)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Bloecker H

MIPS  
Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ). Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
s1 sequence also available.  
This clone (DKFZp434C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZp434C1815"



```

/clone_lib="434 (synonym: htes3)"
/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="vector: pSPori1; Site_1: NotI; Site_2: SalI"

```

```

BASE COUNT      171 a      89 c      138 g      147 t
ORIGIN

```

```

Query Match
Best Local Similarity 100.0%; Pred. No. 1.7e-127;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

14.7% Score 545; DB 10; Length 545;
1028 ccagatagataactttccacagctggaatataaacctgaatccatagaagaactgtg 1087
1 CCAGTATAGTACTTTCCACAGCTGATATTAACCTGATATCAATGAAGAACTGTGG 60
1088 aatttaaaatgttttttccaatatccatcaagaccatctatcaagattctaaagtc 1147
61 AATTAAAAATGTTCTTCAATTAATTCATCAAGACCATCTATCAAGATTCTAAAGGTC 120
1148 tgaatccagaatgaatgctgagagacagctgcctgtgcttccaatgcaatgagag 1207
121 TGAATCTCAGAAATTAAGTCTGGAGACAGATCGCCTTGGTCTCAATGACAGTGG 180
1208 agagtaacgtagtccagctctgcagaggttatatagatccggaatgagcttatactg 1267
181 AGAGTACGAGTACAGCTCTGACAGGTTATGATCCGAGATGATGAGCTTATCATGG 240
1268 tgaatgagatgacatcagagctttaaagtgtgagcatatccagaccatgtgagtg 1327
241 TGGATGGAATGACATCAGAGCTTTAAATGTGGGCAATTAACGACCAATATGGAAGTGG 300
1328 ttgagcaagagctgtttgtctgcggagccacatcagtaacaatlatcaaglatgagag 1387
301 TTATGTCAGAGCCTGTTTGTCTGGGACACACATCACTAATCAATCAATGATGAGAGAG 360
1388 ataatgtgactgtagaagagatgagagagcagcaaggaagaagaatgctatgattta 1447
361 ATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
1448 tcatgtgagtttcccaataaatttaataatctgtgtgagggagaaagagctcaaatgag 1507
421 TCAATGGAGTTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
1508 gaaggcagaagaagagagatgcgaatgtcgtgcttgaatgcgaatgcgaatgcgaatgc 1567
481 GAGGGCAGAAACAGAGATCGCATTCGTCTGCTTGAATTCGAAACCCCAAGATTCGTA 540
1568 ttta 1572
541 TTTTA 545

```

```

RESULT 2
AL520322 943 bp mRNA EST 13-FEB-2001
LOCUS AL520322 LTR_NFL004_NBC2 Homo sapiens cDNA clone CSODB006Y15 5
DEFINITION

```

```

ACCESSION AL520322 GI:12783815
VERSION AL520322.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 943)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France

```

```

FEATURES
source
Location/Qualifiers
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODB006Y15"
/clone_lib="LTR_NFL004_NBC2"
/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"

```

```

/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: filiang@life.com URL:
http://fulllength.invitrogen.com"

```

```

BASE COUNT      253 a      225 c      253 g      211 t      1 others
ORIGIN

```

```

Query Match
Best Local Similarity 65.0%; Pred. No. 3.2e-87;
Matches 386; Conservative 0; Mismatches 314; Indels 1; Gaps 1;

```

```

2795 tagtttactgcaattgcatalatgagatgagcactcgaagaacagctcgtttgctc 2854
6 TATTATTTCAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 65
2855 ctgaatattccaaagccaaatccggggtcgcacatcgttgcctgttgtagaagaac 2914
66 CTGACTATGCGCAAGCCAAATATCAGACGCCACATCATCATCATCATCATCATCATCATCAT 125
2915 caaatatagacccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2974
126 CTTTGTATGACAGCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 185
2975 agttcgaagatcctctctctctctctctctctctctctctctctctctctctctctctct 3034
186 CATTGTGTAAGTTGTAATCAACTATCCACCCGACCGGACATCCAGTCTTCAGGAC 245
3035 taccctcagatattgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3094
246 TGAACCTGAGAGTGAAGAGGCGCAGCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 305
3095 aaagcactctgttcaactctgtcagagactttagacccctgcaagcagcagcagcagcagcag 3154
306 AGACACAGAGTGTCTCAGCTCTGAGCGGTTCTACGACACCCCTTGCGAGGAAAGTGTGTC 365
3155 ttgattgttgatgacaaagaattgaatgacagtggtcgttcccaatlaagacag 3214
366 TTGATGGCAAGAAATTAAGGAGCTGAATGTTAGTGGCTCCGACACACCTGGGCAATGCG 425
3215 ttctcagaagcctgtccttcaactgcagcagcagcagcagcagcagcagcagcagcagcagc 3274
426 TGTCCAGAGACCCATCTCTGTTTACATGACATTTGTTGTAATTTGTTGTAATTTGTTGTA 485
3275 acagcgtgtgtgtcctttagatgagatcaagaagcagcagcagcagcagcagcagcagcagcag 3334
486 ACAGCCGGGTGTGTCTACAGGAAGAGATCTGAGGCGCAAGAGGCAAGCAATACATG 545
3335 ctttattgaaggtctccctgagaataacaacacagaagtgagcagcagcagcagcagcagcagc 3394
546 CTTTCATGAGTCACTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 605
3395 ttctgagcgcagaaacagaagactagctatgcaagggcctctctccaaacccaaac 3454
606 TCTCTGTGGCCCAAAAGCAAGCATTTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
3455 ttctattgttgatgagcagcactcagccctcagatlaatgacagtgagaggttgctcagc 3514

```

DB	666	TTTTGGCTTTTGGATGAAGCCACGTCGTGAGATACAGAAAGTGAAGGTTGTCAG	725
QY	3515	atgcccttgataaagccaggaacgaggagacatgcttagtgcctcaacagcgtctctg	3574
Db	726	AAGCCTCTGGACAAAGCCAGAGAGAGCCGACCTGCATTGTATTGCTCACGCCGCTGCCA	785
QY	3575	caatcagaacgagatttgtagtg-gtcttcgacatagtgaagaatgaagaagaaga	3633
Db	786	CCATCCCAAAAGGACACTTAATAGTGTGTGTTTCAAGATGGCAGAGTCAAGAGCATGTC	845
QY	3634	actcatcaagagcctcctgagaaatcgacacatatatttaagtagtgtaatgcacagtc	3693
Db	846	ACGCATGACGAGCTGCTGCGACAGAAAGCAGCATATTTTCAATGCTGATGTCAGGCT	905
QY	3694	g	3694
Db	906	G	906

  

RESULT	3
LOCUS	AA243820
DEFINITION	AA243820 405 bp mRNA EST 06-AUG-1997
ACCESSION	AA243820
VERSION	AA243820.1
KEYWORDS	GI:1874631
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 405)
TITLE	Hillier L., Allen M., Bowles L., Dubuque T., Geisel G., Jost S., Kucaba T., Lacy M., Le N., Lennon G., Marra M., Martin D., Moore B., Schellenberg K., Steptoe M., Tan F., Theising B., White Y., Wyllie T., Waterston R. and Wilson R.
JOURNAL	WashU-Merck EST Project 1997
COMMENT	Unpublished (1997) Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 1572 Std Error: 0.00 Seq primer: -28ml3 revz ET from Amersham.
FEATURES	Location/Qualifiers
SOURCE	1..405
	/organism="Homo sapiens"
	/db_xref="GDB:556247.3"
	/db_xref="taxon:9606"
	/clone="IMAGE:668506"
	/clone_id="Soares_NHMPU.S1"
	/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
	/lab_host="DH10B"
	/note="Organ: mixed (see below); Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2Nb1M, pregnant uterus NBHPU, and fetal heart NBH19M) were mixed, and ss clones were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT	115 a 81 c 85 g 124 t
ORIGIN	

Query Match	10.4%	Score 385.8	DB 10	Length 405
Best Local Similarity	99.5%	Pred. No. 4.4e-87		
Matches 387	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY 903	tcttcgtcttctcttagtgaatccatagacagatattgcattggagcagcagtcctca 962			
Db 17	TCTCTGTTTTCCTTTAGTGAATCCATAGCAGTATTGCAATGGAGCAGCAGTCCCTCA 76			
QY 963	ctttgaaaccttcgcaatagcccgagagagccttccttcatattttccaggtattgataa 1022			
Db 77	CTTTGAAACCTTCCCAATAGCCGAGGAGCGCTTCCTATATTTTCACAGTTATTGATPA 136			
QY 1023	gaacccacagatagataaactttccacagcctgagatataaactgataccatagaagagac 1082			
Db 137	GAACCCAGATTAATTAATCTTTCACAGCGCTGAGATATAAAGTGAATTCATAGAGAGAAC 196			
QY 1083	tctggaattaaaaaatgttcttccaattatccatcaaacatctatccaagattctgaa 1142			
Db 197	TCTGGAATTTTAAAAATGTTTCTTCATTAATTCATCAAGACCATATTCATCAAGATTCTCAA 256			
QY 1143	aggtctgaatccgaatgaatagctctgagagagacagctgcctctgtgcctgtccaatggcag 1202			
Db 257	AGGCTCTGAATCTCGAATTAATAGCTGAGAGACACACTGCCCTTGGTCGGTCTCAATGGCAG 316			
QY 1203	tgggaaagagtaacggttagtcacagctcttgcagaagttatagatccgagatgagctttat 1262			
Db 317	TGGGAAGAGTACGGGTAGTCCAGCTTCTGCAAGAGGTTATATGATCCGATGATAGCTTTAT 376			
QY 1263	catgtgtgagtgaagaatgacatccagagctt 1291			
Db 377	CATGTGTGATGAGATGACATCAGAGCTT 405			
RESULT 4				
AK020318				
LOCUS	AK020318	872 bp	mRNA	HTC 05-JUL-2001
DEFINITION	Mus musculus adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230106F14, full insert sequence.			
ACCESSION	AK020318			
VERSION	AK020318.1 GI:12860872			
KEYWORDS	CAP trapper.			
SOURCE	Mus musculus (Strain:C57BL/6J) adult male epididymis cDNA to mRNA, clone:1b:RIKEN full-length enriched mouse cDNA library clone:9230106F14.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
AUTHORS	Garniuel,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Methods in enzymology. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
REFERENCE	2. (bases 1 to 872)			
AUTHORS	Garniuel,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome research. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			
REFERENCE	3. (bases 1 to 872)			
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoke,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,T., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Iwawa,M., Ohara,E., Matsubiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuno,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.			
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipillar sequencer			

ch 10.4%; Score 385.8; DB 10; Length 405;  
Similarity 99.5%; Pred. No. 4.4e-87;  
887; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
tttgaacacctgcgaatagccggaggagtgccttcataatttccaaggtattgataa 1022  
tttgaaacctcagaattaagctctgagagacagtcgcttgtcgcttcacatgacg 1202  
gttcgaactcacaagtaaatctgagagacagtcgcttgtcgcttcacatgacg 1202  
agsttcgaattcttcagatttagcttgagagacacagtcgcttgtcgcttcacatgacg 316  
tgggaagagtlacgltagtcacagctctctcagaagttatatgatccgaatga 1262  
gggaagagacagtcgctgtgacagctcttcgcagagagcttatgatcatcgatgacgtttat 376  
ttgtgtgagtgaatagacatcagagctt 1291  
ttgtgtgattgagaattgacatcagagctt 405
```

AK020318 872 bp mRNA HNC 05-JUL-2001  
Mus musculus adult male epididymis cDNA, RIKEN full-length enriched  
library, clone:J9230106F14, full insert sequence.  
AK020318  
AK020318.1 GI:12860872  
CAP trapper.  
Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,  
clone:lib.RIKEN full-length enriched mouse cDNA library  
clone:J9230106F14.

Mus musculus  
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Rodentia: Sclurognathi: Muridae: Murinae: Mus.

1 (bases 1 to 872)  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning  
Methods in enzymology. 303, 19-44 (1999)  
99279253  
10349636  
2 (bases 1 to 872)  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
Itoh,M., Konno,H., Okazaki,K., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome research. 10 (10), 1617-1630 (2000)

20499374  
11042159  
3 (bases 1 to 872)  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsui,K.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Tanabe,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer

JOURNAL MEDLINE  
PUBMED  
AUTHORS

TITLE  
JOURNAL  
REFERENCE

Genome research. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4 (bases 1 to 872)  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5 (bases 1 to 872)

Aachuh,J., Alizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Furumishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiroaka,T., Horii,F., Itoimai,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Koima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sobabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

Direct Submission  
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/,  
Fax:81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAACAAGCCAGACCTCTTTTATTTTTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGACTTCGACGTAAATAATTAATCATCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

Location/Qualifiers  
1..872  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/db\_xref="MGD:MGI:1911565"  
/db\_xref="MGD:MGI:1924956"  
/clone="9230106F14"  
/sex="male"  
/tissue\_type="epididymis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
86..>872  
/note="putative"  
/codon\_start=1  
/protein\_id="BAB32070.1"  
/db\_xref="GI:12860873"  
/translation="MANSERTNGLOETNORYGYPLQEQYKVGNOAVGFIEIFRPNINIDIVMTLTIIGLASINGATVPLMSITVLGILSGCLNGCLVQNPRTKYONCSOTOKLENE DIVLTLYIGIAALIFGVVISFWVTIAROTRRORFPFSILAODISWMDGSIS ICEINTRTGDINKLCDSIGKPIPLMKPONISGEFGIYSTLSKKSKSLSVLTSPLLI MASSALSRLMITSFLSELDAYSRGAQAEBELSSIQTVAFGA"

BASE COUNT  
ORIGIN  
262 a 178 c 187 g 245 t

Query Match 10.2%; Score 378.8; DB 12; Length 872;  
Best Local Similarity 76.0%; Pred.No. 3e-85;

Matches	514; Conservative	0; Mismatches	117; Indels	45; Gaps
Qy	1	ttccgcttgcgtatgacgctgacatcaactcatgatccctggtatctatcctgcatcactg	60	
Db	197	ttccgcttgcgtatgaaatctctggacatttctccctcagctacgtctgggaattttatccatccatc	256	
Qy	61	gtcaatgagacgtcgccttccttaatgscactgagtttaagagaatagatgataacct	120	
Db	257	atgaatgagacacacgcttccctttaaagtccttgaagtccttgaagaaataagatcatTTA	316	
Qy	121	atagatgatatgtctagtcaccaactaacaca-----	150	
Db	317	attaatgatagcctcgtatgtaacaaactaacagaactaaatattacagactgtttccagactcAA	376	
Qy	151	-----taccttcttcaggttgacccctgtattatgttggaaatagtgct	198	
Db	377	GAAMAAGCGAATGAAAGATCATTTGCTGTGACCTCATATTAATTGGAATAGAGACAGCT	436	
Qy	199	gacctgattttgtttacatacagatttcccttgcttgatataactacgacgaagacc	258	
Db	437	gccccatttttggctaaagtaagtaacagatttcccttggctcattactgcaccccggaAAC	496	
Qy	259	aagagattcgaaaaacagtttttcatcagttttgacacagagacatcggctggttgat	318	
Db	497	ACAAGAATCCGAAACAGATTTTTTCATTCAATTTTGGCAAGACATCAGCTGTTGAT	556	
Qy	319	agctgtgacatcgttgaaactaacatcgacatgac---agacattgacaaaatcagtgat	375	
Db	557	GGCAGCTGACATCTGTGAAGACTTAAACCCGCAATGACATGTGATCAACAACTCTGTAT	616	
Qy	376	ggtattgagaataaagattgctctgtttgttttcaaaaatgctactcttttcagttggcctg	435	
Db	617	GGTATTTGAGATTAAGATCCCTCTGATGTGTTTCAAGAACATATCTGGGTTTCTATGTGGCTTG	676	
Qy	436	gcagttggttgytgaagagctggaacatccacctagtgactctatccacgctcctctt	495	
Db	677	GTGATTAAGTTTATATAAAAGCTGGAAGACTCTCCGTTGGTCTGTGTCATCATCTCCTCTC	736	
Qy	496	ataatgacctcaaggcgacagatgctctcagatagtgatcatctcatctgacacagtaagaata	555	
Db	737	ATAAAGGCTTCAATCCGCACTGTGCTCTAGAGATTAATCTCATTTGACCAAGCAAGGCTG	796	
Qy	556	agtgaccttccaaagcctggggcctcttgagagaagaagcttgatcatcatccgaacagtc	615	
Db	797	GATGCTTATTCAAAGCTGGGGCTGTGCTGTAAGAAAGCTTTCATCAATCCAAACAGTC	856	
Qy	616	atagaccttaaggccc	631	
Db	857	ACAGGCTTTGGAGGCC	872	
RESULT	5			
BF692596	559 bp	mRNA	EST	22-DEC-2000
LOCUS	602248949P1	NIH_MGC_62	Homo sapiens	cdna clone IMAGE:4334100 5',
DEFINITION		mRNA sequence.		
ACCESSION	BF692596			
VERSION	BF692596.1	GI:11978004		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
TITLE	1 (bases 1 to 559)			
JOURNAL	NIH-MGC <a href="http://mhc.nci.nih.gov/">http://mhc.nci.nih.gov/</a> .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1998)			
	Contact: Robert Strausberg, Ph.D.			
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>			
	Tissue Procurement: ATCC/DC/DT/DTP			
	CDNA Library Preparation: CLONETECH Laboratories, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			



[illegible]

FEATURES

Location/Qualifiers

High quality sequence stop: 658.

```

/organism="Mus musculus"
/bb_xref="taxon:10090"
/clone="IMAGE:4502552"
/clone_1lb="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/site="Organ: eye; Vector: pcMV-SpOrt6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

```

Query Match	7.4%	Score 275.21	DB 11	Length 726
Best Local Similarity	68.3%	Pred. No. 6e-59		
Matches 396	Conservative 0	Mismatches 183	Indels 1	Gaps
QY 3065	tagcattctgtgaggagcagcgctgtgtggaaaagcactctgttcaacttcgcagagac	3124		
Db 3	TGGCCCTGGTGGGACGACAGTGGCTGGCGGAAAGACACAGTGGTCCAGCGCTCGACGCT	62		
QY 3125	ttatagaccgcgtgcaagagcaagctgcgttttatgtgtgtgatacgcaaaaagatgaatg	3184		
Db 63	TCTPTGACCCCATGCGTGTGATCAGTGCTTTAGATGSGTCAGAAACCAAGAAACCTCAATG	122		
QY 3185	tacagttgcctgcgttccccaatatgaatcgttctcctaagagacgctgcgtcttcaactga	3244		
Db 123	TCCAGTGGCTCCAGACTCAACTGGGCAATGTGTGCCAGGAACCAATTCCTTTGACTGGA	182		
QY 3245	gcattctggaagaatcgcctatgtgtacacagccgtgtgtgtgccattagatgatcatca	3304		

Db	183	GCATCGAGAGACATCGCCCTATGTGAGACACACGCCGGGTGTCGCTCATGTAGATG	242
QY	3305	aagaagccgcgaatgtcagcaaatatccaatccttatttgaagtcctccctygaataaca	3364
Db	243	TGAGGGCGAGCCAAAGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCAAAATATA	302
QY	3365	acacacaagtttgagcttgaagaagagacagcttcttgccgycgcaaaacaagaactagcta	3424
Db	303	ACACAGAGGTGTGGAGACACAAAGGGGACGCGCTCTGTGGGGGCCAAAGACGAGGATTGCCA	362
QY	3425	ttgaaagggcctctctccaaaaacccaaaatttatgttggatgagtcacatcaagccc	3484
Db	363	TCGCCCGAGCCCTCATCGACAGCGCTCGGTCTACTGCTTGATGAAAGCCACGTACGCTC	422
QY	3485	tcgaataatgacagctgagaaggttggttccagcctgcgctgtataaagccagagccggaagga	3544
Db	423	TGGATTACTGGAAGTGAATAAGTTGTGTCCAGGAACACTGTGACAAAGCCACGGGAAGGCCCA	482
QY	3545	catgcttagtggctactcacagagctctctgtcaatcagaaacgcaag-atttataatgtgtt	3603
Db	483	CCTGATTGGCATGTGCTACCGCCTGTCCACCATTCAGAAAGCGCGCACTTGATGTGGTG	542
QY	3604	ctgcacaaatggaagataaagaacaaaggaactatcaag	3643
Db	543	ATTGAGAAAGCGCAAGGTCAAGGAGACGACGCCACCCACAG	582

LOCUS	BG248052	1019 bp	mRNA	EST	13-FEB-2001
DEFINITION	602359987F1 NC1_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5', mRNA sequence.				
ACCESSION	BG248052	1	GI:12757867		
VERSION	EST.				
KEYWORDS	house mouse.				
SOURCE	Mus musculus				
ORGANISM	Eularyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1019)				
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapds@email.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLM10334 row: n column: 21 High quality sequence stop: 650.				

source	
1. 1019	
/organism="Mus musculus"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone_image:4488404"	
/clone_id="NCL-CGAP-Mam1"	
/tissue_type="tumor, biopsy sample"	
/dev_stage="3 months, virgin"	
/lab_host="DH10B"	
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"	
BASE COUNT	276 a 262 c 283 g 196 t 2 others
ORIGIN	
Query Match	7.48; Score 273.2; DB 11; Length 1019;

Best Local Similarity 63.2%; Pred. No. 2,1e-58;  
Matches 437; Conservative 0; Mismatches 253; Indels 2; Gaps 1;

```

OY 2962 gaaggaattagattcgagaagatccttctctatcatcgtccagatgttctc 3021
Db 4 GAAGGAATATGTCATTTAGTGGACTCGTTCACATTCACCCGACCCAGCATCCCA 63
OY 3022 atcccgctgtctatccctcaagatctgacgagaagacgtatcttctgagagc 3081
Db 64 GTGCTTCAGGGGCTGACCTTGAAGTGAAGAGGCCACACCTGGCTGGGGAGC 123
OY 3082 agcgctgtggaagaaagactctgttcaactctgcagagacttaagaccgtgcaa 3141
Db 124 AGTGCCTGGGAGAGACAGACACTGTCCAGCTGTGAGCGCTTCAGACCCCATGGCT 183
OY 3142 ggcacagctgttctgtatgtgtgtatgcacaaagaattgaatgacagtgtcttcc 3201
Db 184 GGATCAGTGTCTGTAGAGGCAAAAGAAATMAAGCAACTGAATGTCCAGTGTCCGAGCA 243
OY 3202 caaataagacatcgttcctcaagagcgtgtctcactgaactgacagatgtctgaacatc 3261
Db 244 CAGCTGGCATTTGTCTCCAGAGAGCCCATCTCTTTGACTGACGATCGCAGAGAACTT 303
OY 3262 gacctatgtgaacaagacgctgtgtgtccatagatgaatcaagaagccgcaatgca 3321
Db 304 GCCTAGGAGAGACAGACGCGGCTGTGTATGAGAGATTTGAGGCGACCAAGAG 363
OY 3322 gcaaatatcattcttatttgaaggtctccctgagaataatacacacaagaattgactg 3381
Db 364 GCCAACATCCACCACTGATCATGCTGCTACCTGTAATTAACACACAGATAGAGAGC 423
OY 3382 aaagagacagcgttcttctgagcgagaaagaagactagctattgcaaggtctcttc 3441
Db 424 AAAGGACCTGACCTGTGGGTGGGAGCAAGCAGCCATCGCATGCGACGCGCTCGTC 483
OY 3442 caaaacaaaatttattgttgaatgagcacttcagccctcagataatgacagtga 3501
Db 484 AGACAGCCTCATTCTTACTGTGACGAGACATCATGCTGTGATACAAAGTGA 543
OY 3502 aagtggttcaagatgccttgaataaagcagagacggaagacatgcctagtgtact 3561
Db 544 AAGGTGTCTCAGGAGAGCCCTGGACAAAGCCAGGAGAGCCGACCTGTGTGATGCT 603
OY 3562 caccagctctcgtcaatcagaagcagatgtgatagtgtt--ctgcaaatgtgaaga 3619
Db 604 CACCCCTGTCACCAATCCAGAGAGCGAGCTGTGATGTGATTCAGAAACCGGAAGC 663
OY 3620 taaagaaacaagaaacatcatcaagagctcctg 3651
Db 664 TCAAGGAGACAGGGCACCCCAACAGCAAGTG 695

```

RESULT 9  
BF796582 944 bp mRNA EST 12-JAN-2001  
LOCUS 602258463p1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4341710 5',  
DEFINITION mRNA sequence.  
ACCESSION BF796582  
VERSION BF796582.1 GI:12101636  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 944)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Louis Staudt, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.lnl.gov

Plate: L1AM9955 row: n column: 15  
High quality sequence start: 9  
High quality sequence stop: 669.  
Location/Qualifiers

#### FEATURES

source

```

1. 944
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4341710"
/clone_id="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

```

BASE COUNT 265 a 241 c 237 g 201 t  
ORIGIN

Query Match 7.3%; Score 269; DB 11; Length 944;

Best Local Similarity 65.4%; Pred. No. 2.4e-57;  
Matches 409; Conservative 0; Mismatches 215; Indels 1; Gaps 1;

```

OY 2860 tattcaaaagcaaatcggggctgcacatcgttctgttctgttgaagaagacaaat 2919
Db 39 TATGCTAAAGCTAAAGCTGTCTGACCCCATTTATCATGCTTTTGAAGAAGAACCTCG 98
OY 2920 atagacagcgcagtcgaagaaggaagaaagcagacacatgtgaagggaattagattc 2979
Db 99 ATTGACAGCTACAGTGAAGAGGGGCTGAAAGCTGTGAATTTGAAGAAATATACATT 158
OY 2980 cgaagatcctcttcttctatcagtcgcagagatgtttcaatctcctgtgtatcc 3039
Db 159 AATGAAAGTGTGTTCAACTATCCACCCAGCAAGAGTGCAGTCTTCAGGGCTGAGC 218
OY 3040 ctcaatattgacgagaaagacagtagcatctgtgtggaagcagcgtgtgtgaagaa 3099
Db 219 CTGGAAGGTGAAGAAAGGCCAGACACTAGCCCTGTGGGAGCAGAGTGGCTGGAAAGC 278
OY 3100 acttctgtcaactctcgcagagacttataagaccctgtgcaagacagtgctgttat 3159
Db 279 ACGGTGTCTCAGCTCTGTGAGGCGTTCAGACCCCTTGGCGGGAACAAGTCTCTCGAT 338
OY 3160 ggtgtgtgtgcaaaagaattgaattagatgtgtgtcgttcccaaatagcaatgttct 3219
Db 339 GGTCAAGAGCAAAAGAAACTCATGTCCAGTGTGCTCAGAGCTCAACTCGAATCGTCT 398
OY 3220 caaagacgtgtcttctcaactgcagcatctgtgagaacatgcctatgtgtgaacagc 3279
Db 399 CAGGAGCCTATCTTATTTGACGACGACATGGCGAAGAAATTTGCTATGAGACAAACAGC 458
OY 3280 cgtgtgtgtccattagatagatcaagaagcgcgaagaaatgtcagcaaatatcatctt 3339
Db 459 CGGTTGTATTCACAGAGATGAATCTGTAGTGCAGCCAAAGCTGCCAACATFACATCTTC 518
OY 3340 attgaagttctccctgtgagaataataacacacaaagttgtgacttgaagagcaagtttct 3399
Db 519 ATCGAGAGCTTACCCCAAAATATGAAGAGAGTGGGAGATGAAGGGAGCTCAGCTCTCA 578
OY 3400 ggcggcgaagaagaagactagctatgtgaaaggtctcttccaaaaacc-aaattt 3458
Db 579 GGAGGTCAAAACAGAGATTTCTATTGCGCCGACGCTCATAGACAAACCCCAATCTCT 638
OY 3459 attgttgatagggcactcagcc 3483
Db 639 CCGTGTGATGAAGCTACATCAGCC 663

```

```

RESULT 10
LOCUS      AV709991      785 bp      mRNA      EST      09-OCT-2000
DEFINITION AV709991 Cu Homo sapiens cDNA clone CUAJ09 5', mRNA sequence.
ACCESSION  AV709991
VERSION     AV709991.1 GI:10728272
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 785)
            Peng,Y., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
            ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,D.,
            Chen,Z. and Han,Z.
            Homo sapiens cDNA Cu clones
            Unpublished (2000)
TITLE      Homo sapiens cDNA Cu clones
JOURNAL    Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES   source
            1..785
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="CUAJ09"
            /clone_1b="Cu"
            /tissue_type="adrenal cortex/adrenal cortex adenoma for Cushing's
            syndrome"
            /dev_stage="adult"
            /note="Vector: pBluescript sk(-)"
BASE COUNT 210 a      180 c      211 g      183 t      1 others
ORIGIN
Query Match      7.1%; Score 262.2; DB 10; Length 785;
Best Local Similarity 64.0%; Pred. No. 1.2e-55;
Matches 458; Conservative 0; Mismatches 254; Indels 4; Gaps 4;
QY 2785 ggcattcagttcttactgcacatgcatatgagctatgacatgcagaaacgctc 2844
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 GGAATTGGAAGAGTATTTTTCAGCTGTGTCTTGTGTGCGCAAGCGCGGCAAGTCA 132
QY 2845 gtttggctccctgaataatccaaagcaaatcgaggctgcgcacatctgttgcctgtg 2904
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 TCATTGTCTCTGACTATGCAAGCCAAAGCCAAATATCAGAGCCCAACATCATCATATT 192
QY 2905 gaagaagaacccaatatagaaagccgacgtcaagaagaagaagaacccaacatgtgaa 2964
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 GAAAGAAACCCCTTGATTGACGACTACGACGAGGAGCCCTAATGCCGAAACACATTGGAA 252
QY 2965 gggattatgagttcgcaagatctcttcttcatcatgctgcacagatgcttctc 3024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 GGAATGTACATTTTGGGAGAGTGTATTCATCATTCACACCGACCGACATCCCAAGT 312
QY 3025 ctccgtggtctaccctcgaatgattgagcgaagaagaacagatagcatattgttgaggagc 3084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 313 CTTCAGGAGACGACCTGAGAGTGAAGAGGCCAGCGCTGCTGTGGCGACAGT 372
QY 3085 ggcgtgtgggaaagacactctgttcaactctgtcagagacttatgaccccggtcaagga 3144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 373 GGGCTGTGGAGACACAGTGTCTCAGCTCTGAGCGGCTTCTAGACCCCTTGGCAGGG 432
QY 3145 caatgctgttattgtgtgtgatgacaaagaatgaaatgacatgacatgctgttccca 3204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433 AAAGTGTCTTGTGATGGCAAGAAATTAAGGACTGATGATGTTCAGTGGCTCCGAGACAC 492
QY 3205 atagcaatcgttctcgaagagcctgtgtcttcaactgcaagcatgtgtgagaacatcgcc 3264

```

```

Db 493 CTGGGCAATCGTGTCCCGAGAGCCCATCTCTTTGACTGCGACATTTGCGAATTTGCC 552
QY 3265 tatgttacaacagc-c-gtgtgtgtccattatgatatcaagaagccgcgaatgcagc 3323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 553 TATGAGACAAACAGCCGGGGGTGTGTACAGAGAAAGATTGTGAGGGCAGCAAGAAGGCC 612
QY 3324 aaatcattcttatttgaaggtccctcctgagaatatacaacacagatgtgactgaa 3383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 613 CAACATACATGCGCTTCATCGAGTCAGTCAGCTT-ATAATATATAGCACTTAACATAGACAA 671
QY 3384 aggaagcagacttcttctgtggtccagaagaagaagactgcatattgcaaggctcttccca 3443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 672 AGGACCTCAGCTCTCTGTGGGCGCCAGAAACACACA-TTGGCCATAGACTGTGCTCCCTGGTGA 730
QY 3444 aaacccaatattattgtgtgagtgagcgaacttcagccctcgaatgacagtg 3499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 731 CAGACTTATAA-TTTGTTTGGATGAAGCCACGCTCAGCTTGTATACAAAGTG 785
RESULT 11
LOCUS      BG587938      780 bp      mRNA      EST      11-APR-2001
DEFINITION BG587938 NHAM Medicago truncatula/Glommus versiforme mixed EST
            library cDNA clone pNHAM-51L19 5' end, mRNA sequence.
ACCESSION  BG587938
VERSION     BG587938.1 GI:13603002
KEYWORDS   EST.
SOURCE     Medicago truncatula/Glommus versiforme mixed EST library.
ORGANISM   Medicago truncatula/Glommus versiforme mixed EST library.
            Eukaryota; mixed EST libraries.
REFERENCE  1 (bases 1 to 780)
            Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
            and Fraser,C.M.
            ESTs from roots of Medicago truncatula after colonization with
            Glommus versiforme, 2001
            Unpublished (2001)
            Contact: Harrison M.J.
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            Noble EST name: N387011e TIGR sequence name: MTDDM70TK More
            information is available at: http://www.medicago.org
            Seq primer: Skm0d (CTA gAA gTA gAT CC).
FEATURES   Location/Qualifiers
            source
            1..780
            /organism="Medicago truncatula/Glommus versiforme mixed EST
            library"
            /cultivar="Medicago truncatula genotype A17"
            /db_xref="taxon:119092"
            /clone="pNHAM-51L19"
            /clone_1b="NHAM"
            /tissue_type="roots colonized with Glommus versiforme"
            /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
            post-inoculation with Glommus versiforme. The library was
            made from a mixture of RNA from each of these stages."
            /lab_host="E. coli strain X10LR"
            /note="Vector: pBluescript SK-; Site.1: EcoRI. Site.2:
            XhoI; cDNA was prepared from polyA+ enriched RNA from
            roots harvested at 10, 17, 22, 31 and 38 days
            post-inoculation with Glommus versiforme. The cDNA was
            directionally ligated into the Unizap XR vector from
            Stratagene and packaged using GigaPack III Gold packaging
            extracts. Plasmids containing cDNA inserts were excised
            from the recombinant lambda-Zap phage using Ex-assist
            helper phage and propagated in X10LR cells."
BASE COUNT 234 a      156 c      184 g      206 t
ORIGIN

```



Query Match 6.8%; Score 253.2; DB 11; Length 780;  
 Best Local Similarity 58.5%; Pred. No. 2.3e-53;  
 Matches 441; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 944 ttgagcagcagtcctccctcattgaacctcgcataagccgagagctgccttcata 1003  
 4 tggggcaatctcaccgacacatggcgcatTTTACAAGGCTAGAGTTGGCGGTGAAGA 63

QY 1004 ttctcaggtttatataagaaccagatagatacttcccaagcttgatataaac 1063  
 64 TTTTCCGGATTAATGATATCCAGCCCTGGATGATGAAGAAACAGTGAATCTGTGAAT 123

QY 1064 cgaatccatagaagaactgtcgaaatttaaaatgtttcttcattatccatcaagac 1123  
 124 TAGAGACACTTCTGACACTTGTGAACCTGAAATGTGACATTTCTTTATCCATCAAGC 183

QY 1124 calcatcaagaatcttgaaaggctcgaatcgaatlaagcttgagagagctgcct 1183  
 184 CTGAAGTTCTGATTCATGATGATTTCTCTTGAGTTCTCTGCGGAAAGACCATAGCTT 243

QY 1184 tggctggtcgaatgagcagtggaagagtaagcagtagtccagcttcgagaggtatag 1243  
 244 TGGTTGGTGGACGGCTCTGCGCAAGACACCTGTTCTCTTATTTGAGAGATCTATG 303

QY 1244 atccgagatgcttatacatggtgagatgagatgacatcagagctttaatgtgcgc 1303  
 304 ATCCAACTTCAGGACAAAGTAATGTTGGATGGCAGTACATTAACCTTGAAGCTTAAT 363

QY 1304 attatcgaccatatttgagtggttagtcgaagagctgtttgttcgagaccacata 1363  
 364 GGTGAGGCAACAAATAGAGTAGTGAGCCAAAGAACCTGCTTGTGTCACACAGATTC 423

QY 1364 gtaacatatacgaatgagagagatgagatgagatgagatgagagagagagca 1423  
 424 GAGAAATATATCTTGGAGAGCCCTGATGACAAACAGGTTGATGAGATGAGAGCTGTA 483

QY 1424 ggaagcaatgctgatatgttatacatgaggttccataaatttaatactgttag 1483  
 484 GGGTTCATATCTCTATTCATTCATCAAACTTCCTGAAGCTTTGAACTCAGGTG 543

QY 1484 gggaaaaagagctcaaatgagtggaagcagaagaagagatcgcaatgtcgtgcct 1543  
 544 GAGAAAGAGACTACAACTTCTGGAGACAAACAAAGATAGCAATAGCAAGGCGAA 603

QY 1544 tggctggaaccccaagatctcgatttagatgaggtcgtcgtccctgagattcaaga 1603  
 604 TGGTAAAAAATCCAGCAATTTCTCTCGATGAGGCAACAGTGTGACTTGAACT 663

QY 1604 gcaagtcagctgttcaagctgacgtgagagcgaagcgaaggtcgactacaatcgttg 1663  
 664 CAGAAAAGCTGTGTGCAAGAGCACTTGACCGGTTCAATGATGGCAGGACAACTCTGTAA 723

QY 1664 tagcacacgaacttctactatctgaagtgcaga 1697  
 724 TTGCTCATGCGCTCTCCACAAATTCGCAAGCTGA 757

RESULT 12  
 BE749379 560 bp mRNA EST 25-APR-2001  
 LOCUS DEFINITION 200194 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BE749379  
 VERSION BE749379.1 GI:10163371  
 KEYWORDS EST.  
 SOURCE COW.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 REFERENCE Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 AUTHORS Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,  
 G.L., Heaton,M.P., Laegreid,W.W., Rehner,G.A., Chitko-Mekorn,C.G.,

Petrea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
 Keefe,J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 Genome Res. 11 (4), 626-630 (2001)  
 21180013  
 MEDLINE  
 COMMENT Contact: Smith rpl  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and all trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACAT  
 BACKWARD: GTTCCGACGACGAG  
 Plate: 108 row: F column: 11  
 Seq primer: ATTAGCTACACTATAG.  
 Location/Qualifiers  
 1..560  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC 4BOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV SPOT6; Site\_1: XbaI; Site\_2: XhoI;  
 library made from pooled tissue from day 20 and day 40  
 embryos."

BASE COUNT 154 a 153 c 146 g 107 t  
 ORIGIN

Query Match 6.8%; Score 252.4; DB 10; Length 560;  
 Best Local Similarity 67.5%; Pred. No. 3.5e-53;  
 Matches 355; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 3171 aaaaagattgaatgtcagtgctccgttcccaataagcaatcgttcccaagagctgt 3230  
 1 AAAGAAGCTCAAGCTCAGTGGCTCAGAGCCCAACTGGATGCTGTGCGAGAGCCGT 60

QY 3231 gctcttcaactgagatgctgagacatcgctatggtgaacagagcgtgtgtgc 3290  
 61 CCGTTTGACTGACGATGTCGACAACTTGGCTTGGGACACAGCCGCTTAAC 120

QY 3291 attagatgatatcaagaagccgcaatgcaagaatataccattcttattatgaagctc 3350  
 121 CATGCCCTGAATGTGTAGGCGACCAAGAGCCAAACATCCATTCATTGACACTT 180

QY 3351 cccctgaaatatcacacaagaatgtgactgaagaagacagacttctcgtgcgcagaa 3410  
 181 GCCCCCAATATGAAACAAAGAGTGGAGTAAAGGAGACTACACTTCCGGGGACAGAA 240

QY 3411 aaaaagactagctattgcaaggtctcttccaaaacaaaattttttgttgatata 3470  
 241 ACAGAGGATGCTATTGCCCCGAGCCCTCATCCGACACCCCGCATCTACTGTGATGA 300

QY 3471 ggcactcagccctcgataatgacagtgagaagtggttcagacatgccttgataaac 3530  
 301 AGCGACGTGACGACTGTGATTAAGTGAAGATTTGTCACAAAGACCCGTGACAAAGC 360

QY 3531 caggacggaagacatgctagtgctacacacaggtctctcgtcaatcagaagcaga 3590  
 361 CCGAGAAAGCCGACCTGCATGCTGATCCCTCACCGCTGTCCACCATCAGAACGACA 420

QY 3591 ttgttagtggttctgcaaatggaagaataagaagacagaagacatcaagagcttct 3650  
 421 CTTGATAGTGTGTAATGCAAAATGGCAGACTCAAGGACAGGACACACCGACGACTGCT 480

QY 3651 gagaatcagacatalatttaagttagtgaaatgacagctcagt 3696  
 481 GGCACGAAAGGACATCTATTTCACCATGTGTCACAGGCTGGG 526



```

RESULT 13
LOCUS BH139685
DEFINITION BH139685 886 bp DNA GSS 07-AUG-2001
ENTNMA477R Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
ACCESSION BH139685
VERSION BH139685.1 GI:15098746
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 886)
AUTHORS Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HML:IMSS sheared DNA library (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: entae@ligr.org
Clones are derived from the Entamoeba histolytica HML:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 18
High quality sequence stop: 851.
Location/Qualifiers
1..886
/organism="Entamoeba histolytica"
/strain="HML:IMSS"
/db_xref="taxon:5759"
/clone_id="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Barelli, Oxford University Press, 1999)."
```

```

BASE COUNT 348 a 105 c 176 g 257 t
ORIGIN
```

```

Query Match 6.8%; Score 252; DB 13; Length 886;
Best Local Similarity 58.5%; Pred. NO. 4.8e-53;
Matches 480; Conservative 0; Mismatches 330; Indels 10; Gaps 2;
```

```

QY 986 gagagcgtccttcattccaggtatgatgaagaaccgcgtatagaacttt 1045
    |||||
DB 12 GAAAGCAGCAAGAAATATGATGATTTATTTGAT-AGAAAAACCTATTGATGTTATA 70
QY 1046 coacagcgtgatataaactgaatccatagaagacgtggaattaaatgtctt 1105
    |||||
DB 71 GTGAAGAAGGCGAAACATTTATGATGTTAAAGCGAATGCAATTTAAAGCATTTGTT 130
QY 1106 tcaattccatcaagaccatcatcaagatctcgaagtcgaatcgaatgaagt 1165
    |||||
DB 131 TCAGATATCCAAACAAGCAGACATCTGCTTGAAGGATTTCAATTCAAAGTAGAAC 190
QY 1166 ctggagagaagtcgcttgctgcgaatgagcgtggagaagatcgcgtatgcagc 1225
    |||||
DB 191 AAGGAAAAACGTTGCAATTAGTAGAGCATCAGATGTGTAAATCAACATCAGTTCAGT 250
```

```

QY 1226 ttctgcagaggtatataatgatccgatgatgcttattcatcgtggtgatgaatgatca 1285
    |||||
DB 251 TGATTGAAAGATTTTATGATGATCCACACATGGAGATGATATTATGATGACATTAATATCA 310
QY 1286 gagctttaaatgctggcgcattatcgagaccatattggaagtggttagtccaagagctgtt 1345
    |||||
DB 311 AAGATTTGAATATTCATTTCTTAAAGAGTCAAAATTTGAATGTAAGCAAGAACCACTAT 370
QY 1346 tttcggagaccaccatcagtaacatatacaagtagt-----acgagatagtga 1396
    |||||
DB 371 TATTTGCTGAAGAGTGTATGATATATTAGAACAGAGATACCTAAAGAGCTTGAAGTAA 430
QY 1397 ctgattgaagatgagagagcaagcaagggaaagcaaatgctgattatcatcagagt 1456
    |||||
DB 431 GTATVGAACAAATTTATGCTGCTCTAAATGCGCAAAATGCACATGACTTATTTCACAAA 490
QY 1457 ttccataataattatcatatcattgtaggggaaagagagtcgaatgagtgagggagga 1516
    |||||
DB 491 TGCCAGAAAGATATACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 550
QY 1517 aacagagatcgcaattgctgctgctttagtcgaagcccaagatctgattagatg 1576
    |||||
DB 551 AACAAAGATTTGCTATTGACAGTGCATTTGATTAAGAAATCCAAAGTGTATTACTGATG 610
QY 1577 agctacgtctgccttgatcagaagcaagcaagtcagctgctcagctgagagaag 1636
    |||||
DB 611 AAGCTACATCAGCACTTGATTCAGAAAGTGAAGATTTGCAAGATGCACTTGACAAAG 670
QY 1637 cgagcaaaagtcgagcatcaatcgtgtagcagaccgacttctactatctgaagtgag 1696
    |||||
DB 671 CAGCAAAAGGAAGAACCAATTTAATGCAATAGATTATTCATTAATTAATTAATGAG 730
QY 1697 attgatgtgaccctaaagatggaatgctgcgagcgagaaagagacacatgctgaactaa 1756
    |||||
DB 731 ATCAATATGCTTATATATGAGAGCAAGCAATTTGCAAGAAAGCAACATCAAGATTA 790
QY 1757 tggcaaaagaggtcatattatcactgattgattgacaa 1796
    |||||
DB 791 TAGATTGAAGATTTATATATACACTTGTCTATGCAACA 830

RESULT 14
LOCUS BF969667
DEFINITION BF969667 981 bp mRNA EST 22-JAN-2001
ENTNMA477R Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
ACCESSION BF969667
VERSION BF969667.1 GI:12336882
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 981)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LRAM10000 row: 1 column: 11
High quality sequence stop: 670.
Location/Qualifiers
1..981
/organism="Homo sapiens"
/db_xref="taxon:9606"
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 23:04:06 ; Search time 886.72 Seconds  
(without alignments)  
3579.279 Million cell updates/sec

Title: US-09-873-409-13  
Perfect score: 3702  
Sequence: 1 ttccgcttgcgtatgact.....atgcacagtcagtcagtga 3702

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 segs, 428662619 residues

Total number of hits satisfying chosen parameters: 1661242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N\_Geneseq\_1101:\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1234.4	33.3	4279	22	AAD03504
2	1234.4	33.3	4279	22	AAD03505
3	1234.4	33.3	4279	22	AAD03506
4	1234.8	33.3	4279	22	AAD03488
5	1229.6	33.2	4317	22	AAD03489
6	1225.2	33.1	4349	22	AAH57442
7	1225.2	33.1	4646	15	AAQ72872
8	1225.2	33.1	4646	21	AAZ94738
9	1225.2	33.1	6505	17	AAI13394
10	1223.8	33.1	4669	14	AAO52726
11	1223.6	33.1	3860	21	AAZ49332

12	1223.6	33.1	4669	8	AAH70752
13	1223.6	33.1	8630	21	AAZ24042
14	1222	33.0	8630	21	AAZ49333
15	1220.4	33.0	8630	21	AAZ24041
16	1218.8	32.9	4378	11	AAQ04522
17	1217.2	32.9	4669	19	AAV32645
18	1215.6	32.8	4186	22	AAH6127
19	1215.6	32.8	4195	22	AAH6128
20	1214.2	32.8	4264	19	AAV6533
21	1212.6	32.8	4264	19	AAV6534
22	1209.2	32.7	3988	21	AAZ88973
23	1185.2	32.0	3924	21	AAZ94742
24	1185.2	32.0	3924	21	AAZ88974
25	1125.8	30.4	4425	21	AAZ52048
26	1125.4	30.4	4189	21	AAZ49334
27	1125.4	30.4	4313	14	AAQ38950
28	1124.2	30.4	4369	21	AAZ52047
29	1121	30.3	4233	22	AAZ90198
30	1121	30.3	4233	22	AAZ27498
31	1113.2	30.1	4788	21	AAZ49335
32	856	23.1	2726	15	AAO70907
33	856	23.1	2726	15	AAO70916
34	856	23.1	2726	18	AAZ43322
35	773	20.9	4176	21	AAZ94744
36	509.6	13.8	4175	20	AAV63392
37	436.4	11.8	4051	21	AAAO9019
38	433.4	11.7	4102	21	AAAC45942
39	426.6	11.5	4002	20	AAH85823
40	386.6	10.4	4047	20	AAH60201
41	383.8	10.4	2698	20	AAV63394
42	349.6	9.4	3512	19	AAV63395
43	331.4	8.7	4224	19	AAV42347
44	322.6	8.7	8777	22	AAH83399
45	322.6	8.7	8777	22	AAH83400

## ALIGNMENTS

RESULT 1	
AAO03504	standard; cDNA; 4279 BP.
ID	AAO03504;
AC	13-JUN-2001 (first entry)
DT	Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX	Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1; drug bioavailability; transgenic animal; genetic model; ss.
KW	Canis familiaris.
OS	Canis familiaris.
XX	Key
PH	Location/Qualifiers
FT	CDS
FT	/*tag= a
FT	/product= "Dog P-glycoprotein (PGP) allelic variant (Genotype A) protein"
FT	replace (607, C)
FT	/*tag= b
XX	allele
PN	WO20013540-A2.
XX	05-APR-2001.
PD	28-SEP-2000; 2000MO-US26767.
XX	28-SEP-1999; 9905-0156510.
PF	(GENT-) GENTEST CORP.
XX	28-SEP-1999; 9905-0156510.
PR	Stocker PJ, Steimel-crespi DT, Crespi CL, Relf TC, Patten CT;
XX	
PI	

Sequence of human  
Retroviral vector  
Human G185V mutant  
Retroviral M4 mdr-  
Multidrug Resistant  
Human P-glycoprotein  
Cynomolgous monkey  
Mutated human P-gly  
Mutated human P-gly  
Human MDR-1 DNA.  
Human ATP binding  
Human MDR-3 DNA.  
Rat multidrug resis  
Rat multidrug resis  
Rat mdr1b2 multidr  
Murine multidrug r  
Multidrug-resistant  
Multidrug-resistant  
Human ATP binding  
Human ATP binding  
H. contortus PGP-A  
Arabidopsis thaliana  
cDNA encoding mult  
cDNA encoding the  
H. contortus PGP-A  
H. contortus PGP-O  
cDNA encoding a mu  
P. chrysogenum ABC  
P. chrysogenum ABC

```

XX WPI: 2001-235373/24.
DR P-PSDB; AAE00308.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
XX Claim 9: Page 85-90; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) allelic variant
CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC exporting small molecules across the cell membrane. This enzyme
CC is a member of the ABC transporter family.
XX
XX Sequence 4279 BP: 1295 A; 833 C; 1008 G; 1143 T; 0 other;
XX
Query Match 33.3%; Score 1234.4; DB 22; Length 4279;
Best Local Similarity 61.2%; Pred. No. 0;
Matches 2190; Conservative 0; Mismatches 1256; Indels 132; Gaps 7;
XX
QY 162 caggtgacccgtattatgttggaatggtgtgctgcttgattttgtatcacaca 221
DB || || || || || || || || || || || || || || || || || || || ||
QY 222 gattccctgtgatataactgcagacagacagcaacaagatctcgaaacagctttt 281
DB | || || || || || || || || || || || || || || || || || || ||
DB 418 ggttccatctgtgctctgcagcagaagaacagatactccaattgaacaaacatttt 477
QY 282 tcaatcagtttgagcagacacatcgctgtgttgatagctgtgacatcggtgaactaa 341
DB || || || || || || || || || || || || || || || || || || || ||
DB 478 tcatgcatatcatgcagcagagatgtgtgttgatgcatgacgttggggagcttaa 537
QY 342 cactcgatgacagacattg---acaaatcagtgatgatttggaataagattgctct 398
DB || || || || || || || || || || || || || || || || || || || ||
DB 538 caccggctcacagacagatgtctccaataatcaatgaaggaattggcacaaattggaat 597
QY 399 gttgttccaacaacatgtctacttctcatgtgctgcgcagttggttggtaagggtcg 458
DB || || || || || || || || || || || || || || || || || || || ||
DB 598 gtctcttcaatcaatagaacaattttccacgggtttatagtggtgtttacaagtgttg 657
QY 459 gaaactaaccttagtgactctatccacgctctcctctaataatggtccagcgacgatg 518
DB || || || || || || || || || || || || || || || || || || || ||
DB 658 gaaactaaccttgtgattttgcccacagccctgtctcttgacttccaagccgcatctg 717
QY 519 ttctaggaatgtcatctcatgacagtaaggaattgaagtcctattccaaggtggg 578
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 718 ggcagaagatcatcttcatcttacttgataaagaactcttgccatgcaaaagctgagc 777
QY 579 tgttgacgaagaagctctgtcatcatccagacagatcatagccttggggccggcgaa 638
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 778 agtagcggaagaagctcttagcagcaatcagacagactgtatggtccttggagcaaaagaa 837
QY 639 agaacttcaaaagtccttccctttaataataacaagatagtctgtttatttcccca 698
DB || || || || || || || || || || || || || || || || || || || ||
DB 838 agaac----- 842
QY 699 gtggtactaagtgtgtctgttnttgttaaggtatatacagagatttcaaaagtacaa 758
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

DB 843 -----tgaagaagtacacaacaaatttagaagaagctaa 876
QY 759 ggaatttggcataaaaggactatagcttcaaaagtgctctgtgctgtgactctt 818
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 877 aggaatttggataaagaagaactatccacgcgcacacattctatgtggcgtcttatt 936
QY 819 tatgaatggactatagactgtcttggttgttgtaacctcttgattcttaagtaga 878
DB || || || || || || || || || || || || || || || || || || || ||
DB 937 gatctatgcatcatatgtctgtctgtgtatgggaacctctgtctccacg--- 993
QY 879 acctggtatcacatcgagactgtctgtcttcttctttagtgtaatccatagcgatla 938
DB || || || || || || || || || || || || || || || || || || || ||
DB 994 -----tgaattatctatgtgaaagaagtactcaactgtctcttcttgatatttgggctt 1050
QY 939 ttgatgtgagcagcagtcctccatcttgaaccttcgcaatagcccgagagctgctt 998
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1051 tagtatgtgacagcagctccccaagcattggaagcatttgcacaagcagagagactta 1110
QY 999 tcatatttccaggttctgtataagaaccagtatagataactttccacagctgata 1058
DB || || || || || || || || || || || || || || || || || || || ||
DB 1111 tgaatcttcaagataatttgcacataaacaagcatgtacagcattcgagaagtgcaca 1170
QY 1059 taaacctgaaatccatagaaaggaactgtggaatttaaaatggttcttcaattatccatc 1118
DB || || || || || || || || || || || || || || || || || || || ||
DB 1171 taaacagataaatatgaaggaatttggaaatccaataaagtctcaactgcttaaccttc 1230
QY 1119 aagaccatcatcaagaattctgaaaggtctgaaatcagaattaagtctgagagagact 1178
DB || || || || || || || || || || || || || || || || || || || ||
DB 1231 tcgaaagaagtttaagttcttaaaaggtctcaacctgaaagtttaaggttggcagagact 1290
QY 1179 cgcttgtgtcgttccatgtgcagtggaagagagagctgtatccagcttctgcagaggt 1228
DB || || || || || || || || || || || || || || || || || || || ||
DB 1291 ggcgtgtgttgggaacagtgctgcggaagagagacacccgtgcagctgacgagagct 1350
QY 1239 atatgacgcgagatgatgtcttatacatgtgagatgagaatgcacagcatttaagt 1298
DB || || || || || || || || || || || || || || || || || || || ||
DB 1351 ctatgaccccaacagatgcatgtctgtatgtatgagcagagacattgagccataaatgt 1410
QY 1299 gcggcattatcgagaacatatgtgagtgtagtcaagaagcgtgttctgtcggagcac 1358
DB || || || || || || || || || || || || || || || || || || || ||
DB 1411 aagcacttctcgggaataattctgtgtgtgagtcagaagacgttgtgttggccacac 1470
QY 1359 catcagtaacaatcatcaagratgagcagatgtatgtagcgtatgaagaagctgagagac 1418
DB || || || || || || || || || || || || || || || || || || || ||
DB 1471 gatagcgtgaanaaactctgcatactgcgcgaaatgltcacatgagatgtgagaagc 1530
QY 1419 agcagaaggaagcaaatgcgtatgtatcttcaatgagtttccatataatataaat 1478
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1531 tgttaaggaagccaatgctcatgtatttatacatgaactactataatttgacactct 1590
QY 1479 gtlaggggaaaaaggagctcaaatgagtgagggcagagaaacagagatcgcaattgctg 1538
DB || || || || || || || || || || || || || || || || || || || ||
DB 1591 gtttggagagagagggccacgctgagtggtgagcagaagaaggaatcgcatctgtcgt 1650
QY 1539 tgctttagttcgaaaccccaagattctgtattgtagaggtcagctccctcgagctc 1598
DB || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1651 ggcctgtgttcgcaaccccaagattctctgtcgtgagatgagcaagctccagtcggcac 1710
QY 1599 agaaagaacgtcagctgtcaagctgacttgagagaagcgagcaaaagtcggaactaaat 1658
DB || || || || || || || || || || || || || || || || || || || ||
DB 1711 tgaagtggaagcagtggttccagtggtgccttgatagaagccgaaagcggaactacat 1770
QY 1659 cgtgtgacacacccgacttctactattcagaatgtcagattgtattgacctcaagga 1718
DB || || || || || || || || || || || || || || || || || || || ||
DB 1771 tgtatagctcatcagttgtgtctacagtgtgaagccgatttctatctgtgtttgtga 1830
QY 1719 tggatgtctgcggaagaagagacacatgtcgaactaattgccaagaaggtctatata 1778
DB || || || || || || || || || || || || || || || || || || || ||
DB 1831 tggagtcatttgtgagaagaagaatcatcagtaacatcgaatgaagaagaggtcttact 1890
QY 1779 ttcaactgtgtgcacaggaatataaanaagcgtgata--acaaatgtgagtaacat 1837
DB || || || || || || || || || || || || || || || || || || || ||
DB 1891 caaactgtctcaatgcagaaagagaaatgaaatgtgagttagaanaaatgacctgtgtga 1950

```

1838 attcactgaaagaagaagccaactcactctctgactctgtgaagag-----cat 1889  
1951 atccaaagctgaagctgctgctctggaatgtctccaaagaattcagggtccagtttat 2010  
1840 caagtcagactcattctgcaagcgtcgtgaagatccccaattcaagagataagctt-- 1947  
2011 aaaaagaagatcactcgcagagatatacatgacacacagccagacagaagaactctgg 2070  
1948 -----cctgaagctcctcattataaaatttaagtt 1979  
2071 tacaaaagagacttgatgaaatgtaacgtccagcttctctctgtgagagattctgaact 2130  
1980 aaaaagcctgaatggtcctttgtgtctgtgagacattgtctctgttctaagtgaaac 2039  
2131 gaactcaactgaatggtccttattgtgtgtatatttctgtatatttataataaaggaag 2190  
2040 tgtcatccagatatttccatcatccttctgcaaaaattaaaccaatggtt--ggaataa 2096  
2191 cctgcaaccagcattttcaataataatttcaagataataaggatcttaccggatgata 2250  
2097 tgataaaacacatlaaagcatgagcagaatattatccatgatatctgtcatatttgag 2156  
2251 ggaacctgaacaaacagacagaatagtaacatgttctctgtattgttctagtctctgg 2310  
2157 tgtatttgccttgcagttatttccatgcaaggatatttaccgacagacaggggaat 2216  
2311 aattatcttcttattatatttctcccaaggtcttcaattcgttcaagacgttggagat 2370  
2217 tttaacgattgatatgaacacttgccttcaagcctatgtatatccagatattgcctg 2276  
2371 cctcactaagcgtctcgtatcatagtttccagatccatgctggaacacggatgtcaactg 2430  
2277 gtttgatgaaaaagaaacagacacagagcgtctgcaacaacatlaagccatagatatagc 2336  
2421 gtttgatgacccataaaacacacacagcagcatttgcaaacacagcgttgcacatgacgcgc 2490  
2337 acaaatlaaagagcaacaggttccagatattggtccttcaacaacaatgtcacaactacat 2396  
2491 tcaagtttaaaagggtctaaagttccaggtctgtctgtcttccccaagatalatgcaactat 2550  
2397 ggagacttcaattatcattccttataatagatgagatgagatgagatccctgattcctgag 2456  
2551 tgggacagcgtattatatacttaactatctatgtgtgtgcaatlaaacacttctcttgc 2610  
2457 taatgtccagtaacttgcgtgacaggaatgattgaaacccgacgaacatgactgtttgc 2516  
2611 aattgaccatcatgtgaatagcagaggtgtgtgaaatgtaaatgtgtctgcgacaagc 2670  
2517 caacaagaataagcaagaacttaagcattgctggaagaagatagaactgaagcttggagaa 2576  
2671 actgaataataagaagaagctagaagagcgtgggaagattgtctaaagaagccatcgaa 2730  
2577 tatcgtaactatgtatcatttaacaagaagaaagccttcgagcaaatgtatgaagaat 2636  
2731 ctccgaactctgttcttcttgactcgggagcagaagtttgaaatacgtatgacgcagag 2790  
2637 gcttcaagactcaacaagaataacctcgaagaagaagacagattatgtgaagctgtatgc 2696  
2791 ttgtgaagtacacatacagaactccttggaggaagcacacactcctcggtctcatcttc 2850  
2637 attcaagcattgccttataataatttgcctatgacagaggttgcatttgagactattt 2756  
2851 tatcaccaggaatgatagtatttctcctatgtcgtgtctcggtttgggcctactt 2910  
2757 aattcaagctggaagatgaccccgagagcattgtcatagtatttactgcatgcatata 2816  
2911 ggttgcaaaatgattcagaacttccagatgtcttcttgatctccagcattgtctt 2970  
2817 tggagctatggtccatcggaanaacgctcgttttgcctcctgaataatactcaagccaaat 2876  
2971 tggtagccatggtgaggtggcaggtcatgtcattgtctcctgactgacatgccaagccaaagt 3030

2877 ggaggctgcgcatctgttctgttctgttggaaaaagaacccaatatagacagcccgacta 2936  
3031 atcagcagcccccgcgtcatcatgatacttgaataaaagccctctgattgacagctacagccc 3090  
2937 agaaaggaaaaagccagacacatgtgaaagggaatttagatttgcgaagaactcttctt 2996  
3091 tcaagcgtcccaagccaaatcacgttggaaagaaatgtgacatttaagtaggtctgttcaa 3150  
2997 ctatccatgtgcgccagatgatttcaactcctcgtgtgcttataccctcagatattgagcagag 3056  
3151 ctatcccatctgacacagacatcccgctgtctcaggggtcgagacctcgaagggtgaagaagg 3210  
3057 aagaagcagtagcattgtgggagcagcgtcgttgggaagaaacactcttgatcaacttct 3116  
3211 ccagacgcgtgcctctcgttagcagcagtgctgtggaagacacagttgttcaagctcct 3270  
3117 gcagagactttagaccccgctgcaaggagaatgtcgttctgtatgtgtgattgataaaga 3176  
3271 agagcgtctctatgaccccttgcgtgtcagtgctaatgtatgtgcaagaagataaagca 3330  
3177 attgaatgacagtggtcgtctcgttcccaatlaagcaatcgttccctcaagacgtgtctctt 3236  
3331 cctgaatgtccagtgctcgcgagcacacctggcagcattgtctcagagacccatcctgtt 3390  
3237 caactgcagcattgtcgtgagaacatgcgtatgttgcaacaagccgtgtgtgcatatga 3296  
3391 tgactgcagcattgtcgcgagaacatgtcctatggaagacaacagcgtgtcgtatcaatga 3450  
3297 tgagatcaagaagacccgcaatgcaagaatatccatcttcttattgaaagttccctga 3356  
3451 agagattatgacggtgcagcaaggaaggccaacataacccatctcatctgagacactccctga 3510  
3357 gaatacaacacacaagaattgtgactgaaaggagacagcttctgtgcgcgcaagaacaaag 3416  
3511 gaaatacaacacacagatgagtagagacaaaggaaacccgctctgtgtgcagagaacagcg 3570  
3417 actagctattgaaagggtcctctccaaaacccaatatttatgttggatgagggccac 3476  
3571 catgccaagctcgcgcctcctgttgaagacagccctcatatttgccttggatgaaagctac 3630  
3477 ttcaagccctcgaataatgaaagctgagaagtggttctcagcatgtcccttgaataagccagagc 3536  
3631 atcagctcgtgatacgaagaagtgaaagggtgttccaaagaagcccttgacaaagccagaga 3690  
3537 gggaagagcattgcctagtggtcactacagcgtctctgtcaattcgaagcagattgat 3596  
3691 aggcgcgacccctgattgtatgcgtccaccgcgtgtgtccaccatccagatgcagattta 3750  
3597 agtgttctgcacaatggaagaataaagaagaagaactcatcaagaagctccttgagaa 3656  
3751 agtgtgttctcagaatggaagtcagaagtcagatgacacatacacaacgtgtgtgcga 3810  
3657 tcgagacataatttaagttagttgattgacacagtcag 3694  
3811 gaaagcactatatttccatgtgtcaggttccagcgtg 3848

RESULT 2  
AAD03505  
ID AAD03505 standard: cDNA: 4279 BP.  
XX  
AC AAD03505;  
XX  
DR 13-JUN-2001 (first entry)  
XX  
DE  
XX  
XX Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.  
XX  
XX Dog: P-glycoprotein allelic variant; PGP: multidrug transporter;  
KW MDR1; drug bioavailability; transgenic animal; genetic model; ss.  
XX  
OS Canis familiaris.  
XX  
FH Key Location/Qualifiers

```

FT CDS 17..3862
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) allelic variant
FT (Genotype B) protein"
FT replace (91, T)
FT /tag= b
FT /tag= c
FT allele
FT /tag= c
FT WO200123540-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000MO-US26767.
XX
XX 28-SEP-1999; 99US-0156510.
XX
XX (GENT-) GENTEST CORP.
XX
XX Stocker PJ, Steinmel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
XX WPI: 2001-235373/24.
XX
XX P-PSDB; AAE00309.
XX
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
XX Claim 9; Page 93-99; 111pp; English.
XX
XX The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are
XX useful for determining the bioavailability of drugs and for
XX screening PGP inhibitors. They are useful for the diagnosis and
XX treatment of conditions characterised by PGP activity, by
XX reducing or increasing PGP activity in a cell. PGP nucleic acids
XX are used as oligonucleotide probes. Complements of PGP nucleic
XX acids are useful as antisense oligonucleotides, to induce a PGP
XX 'knockout' phenotype. They are used to prepare a non-human
XX transgenic animal, which are valuable as genetic models for
XX human diseases.
XX The present sequence is dog P-glycoprotein (PGP) allelic variant
XX (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
XX exporting small molecules across the cell membrane. This enzyme
XX is a member of the ABC transporter family.
XX
XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;
SQ

```

```

Query Match 33.3%; Score 1234.4; DB 22; Length 4279;
Best Local Similarity 61.2%; Pred. No. 0;
Matches 2199; Conservative 0; Mismatches 1256; Indels 132; Gaps 7;

```

```

Db 598 gtcttcaatacaatagacaattttcacccggtttatagtgagggttaacagtggtg 657
Qy 459 gaactacacctagtgactctatccacgctctctcttaataatggtctcaacgagcatg 518
Db 658 gaagctaaccttgatgttttggccacacgacctgttcttggacttgcacgacatgct 717
Qy 519 ttctagatgtctcatcttgaccagtaaggaattagtgctctattccaaagctggggc 578
Db 718 ggcacaagatactatctcaattactatgataaagaaccttggcctatgacaagctggagc 777
Qy 579 tgtgacagaagaagtcctgtcatccacacgcagatcatagactttagggccagagaa 638
Db 778 agtagcgaagaagtccttagcaacatcagacatgctgattgctttagagacaaagaa 837
Qy 639 agaacctcaaaagtgcttctctttaaataatacaagatagcttggtttatttcccca 698
Db 838 agaac----- 842
Qy 699 gtggtactaagtgtgttctgtntttgttaaggatatcacagatccaaagtcaaa 758
Db 843 -----tgaaaggtcacacaacaataattagaagaagctaa 876
Qy 759 ggaatttggcataaaagagactatagcttcaaaagtgtctctgtgtgtgtactctt 818
Db 877 aggaatttgggataaagaagatcatccacggccacacattctatttggcgttcttatt 936
Qy 819 tatgaatgaaacctatgactgcttctttagtargaaacctcttgattcttaatgaga 878
Db 937 gatcatgacatcatgctctgtcttctgttatggacacctcttggctccctccag-- 993
Qy 879 accggtatcatccatcggaactgtctgtctgttcttctttagtgaatccatacagta 938
Db 994 ---tgaatattctatggacaagaatcatccactgtcttcttctgtatatttgggctt 1050
Qy 939 ttgatttggagacagatgccctccactttgaaaccttgcacatagccagagagctgtc 998
Db 1051 tagatttggagacagatcccccagacatggaagcatttgcacagagaagagagctta 1110
Qy 999 tcataatttccaggttattgataagaacccaagatagataacttttccacagctgata 1058
Db 1111 tgaattcttcaagataattgacaataaaccagacatgtacagcatctgaaagagtgaca 1170
Qy 1059 taaacttgaatccatagagaaggaactgtgaaattttaaattgttcttcaattatccatc 1118
Db 1171 taaacaagataataatgaaggaaatttggaaatttcaaaaattgttcaattcagttaccctc 1230
Qy 1119 aagaccatctacaaagattctgaaagtgctgaatctcagaatttaagctgagagacagt 1178
Db 1231 tcgaaagaagtttaagatccttaagaaggtctcaacctggaaggttcagagtggcagacagt 1290
Qy 1179 cgcttggctggtctcaatgagcagtgaggagatgcgttagtccagcttctgcagaggt 1238
Db 1291 ggcgttggttggacaacatgtgtctgcgggaagagcagcgccgctgagatgcaagagct 1350
Qy 1239 atatgactcgatgctgttcatcatggttgatgagatgacatcagacagctttaaagct 1298
Db 1351 ctatgaccccaacagatggcagatgctctgatatgataagagacatgagaccataaagct 1410
Qy 1299 gcgcattatcgagaccatattgagatgttgaagaaagcctgttcttgcggagcaac 1358
Db 1411 aaggaatcttcgggaatattactagctgtgtgtgtagcagagcctgtgtgttgcacacac 1470
Qy 1359 catcagtaacaatatcaaglatgacagagatgatagtgactgtatgaagagatgagaagc 1418
Db 1471 gatagtcgaacaacatctgtatgcccgcggaataattgcacacagatgtagatgagaagac 1530
Qy 1419 agcaagggaagaagaatgtagatgatttatactatgagatttccctaataatttaacatt 1478
Db 1531 tgttaaggaagaagcatgctcatgatttatactatcagaaactaactaataatttgacacct 1590
Qy 1479 ggttagggaaaagaagagctcaaatgagttagagggagaaacagagagatcgcattgtctg 1538
Db 1591 ggttgagagagaaggggcccagctgagtggtgtgacagaaacagagaatcgcattgtctg 1650

```

OY	1539	tgccttagtcgaaccccaagatctgattttagatgagctgctgcctgagtc	1538
OY	1540		
Db	1651	ggcccttgctgcgaaccccaagattcttctgtctgataagcaagctcgagac	1710
OY	1599	agagaagaatcgagctgcttcaaacctctgacatgtgagaagcgaaacaggttgactaacat	1658
Db	1711		
OY	1719	tggaaatgctgscgagaagaagcgacatctgtaactaatgccaacgaagtcatacta	1778
Db	1831	tggaatcattctgagaaaagaagaatcagaatgaactcatgaaagagaagagcatctact	1890
OY	1779	ttcaactgtgatctcacaagataataaaaagctgata-acaga.tgagtcgaatgacat	1837
Db	1891	caaatctgcacaatgcaagaagaaggaatgaatgtgattgagaagaatgcacgtgga	1950
OY	1838	attctactgaagaagaagccaactcactctctctgactctgtgagaag-----cat	1889
Db	1951	atccaaagaatgaaatgagctgtgtaaaatgctccaaagaattcaaggtccagtttat	2010
OY	1890	caagctcgaactcatttgaaagcgctgaggaatccaccccaatcaagataaggtct--	1947
Db	2011	aaaaagaagatcaactcgccagaggtacacatgcacacaaagcgcaagcagaagaagcttg	2070
OY	1948	-----cctgaagctctctcatnaaaaattttaagtt	1979
Db	2071	tacaaaagatgacttgatgagatgtactccagattccctctgtgagattctgaagct	2130
OY	1980	aaacaagcctgaaatgaccttctgtggtctctgggacattgctctgttcaaatggaac	2039
Db	2131		
OY	2040	gaactcaactgaaatgaccttattctgtgtgtgattatttctgattataaaacggagcg	2190
Db	2191	tgatcatcagatatttccatcatctcttgcacaaattaaacaaatgctt--ggaataaa	2096
OY	2217	tttaacgatagataaagaacacttggccttcaagcagatgtatatacaggaatgtacct	2276
Db	2371	cctcaactaaagcgcttcgataacatgtttccagatccatgctgagacagatgtcagctg	2430
OY	2277	gtttgataaagaagaagaacgacgagaggtctgcaacaataatagccatagatatgc	2336
Db	2431	gttttgatgaccttaaaaaacccaactgagatctgtaaacacaggtctgccaatgagtgcgc	2490
OY	2337	acaaatccaagaggaacacaggttccagagtttgctgcttcaacacaaaatgcaactaacat	2396
Db	2491	tcaagttaaaagggcctataggttccaggtctgtgcattaccagaataatagaacaact	2550
OY	2397	gggaacttcagttacatttctcttatataatgataatggagatgacatcctgattctgag	2456
Db	2551	tgaggaacgaggaattatatacttaactcaatcatgattgtgccaatnaaacatttaactttagc	2610
OY	2457	tattgtcccaagtaattgcgcgtgacgaagatgtttgaaacccgacgaatgagatgagttgc	2516
Db	2611	aattgtatcccatcatctgcaatgacgagagttgttgaagaatagtaaatgtgtctgcgaacgc	2670
OY	2517	caacaaagataagcaagaacttaagacatgctgagaaagatagcaactgaaagctttggagaa	2576
Db	2671	actgaaagaataaagaagaagctaaagaagatctggaaagattgtctaaagaagacatcgaaaa	2730

OY	2577	tatcgtactatagtggtcatttaacaaagggaagaaagccttcgagcaaatgtagtagagat	2636
Db	2731	cttcgcgaactgtgttcttccttgactcgggagcagaagatttgatacatgtagtgcacaga	2790
OY	2637	gcttcagactcaacacagaagaataccctccgagaagaagacagatattatggaagctgtgac	2696
Db	2791	tttgcgaagttacatacacagaataacctcttgagaagaacacacactccgggtctcatcttc	2850
OY	2697	attcagacatgcctcttatatatcttgcctatgcagcagaaggtcttcatttggagccattt	2756
Db	2851	tatcacccagcgaatgtagttatcttccatgctgttcctcggtttgcggtttgacctactt	2910
OY	2757	aattcaagcttgcagcgaatgagcccccagagggcatggtcatagttttactgtcaattgcata	2816
Db	2911	gggtgcgaatgagttcatatgactatgcagacttcagagatgtctcttgttatctccagctattgctt	2970
OY	2817	tggagctatggcccatcgggaaaaagcgtcttgggtccctgtagattatccaaagcccaatc	2876
Db	2971	tgggtgcacatgvcagttggggcaggtcagttcaatttcgtccctgcagctatgcccagaagtc	3030
OY	2877	gggggctgcagatcgtgttgcctgtgttgaaagaagaacaatatagaaagccgcagttca	2936
Db	3031	atccgcagcccaagctcatcatgatactatgaaaaagccctccctgatgtgaaagtaagcccc	3090
OY	2937	agaaagggaanaagccagacacacatgtgaaaggaaattagaagtttcgagaagctctcttc	2996
Db	3091	tcacgctccatcgaagcaaatcgtttgaaagaaatgtagcattaagaaggtcgtgtttca	3150
OY	2997	ctatccatggtgcgccaaatgttttcatcctccgtgtgcttatccctccagttatggacggg	3056
Db	3151	ctatcccaactcagacaaacacatcccgctgtccacgggggcgtgaagcctccaggtgaagaaggg	3210
OY	3057	aaagacagtagcatcttggggggagcagcgtcttggaagaagaacatctggttcaactct	3116
Db	3211	ccgaagcgtctgcgcctcgtgaagtgaagtcaggtgctgttggaagagacacagttgttcagctct	3270
OY	3117	gcgaagacattatgaccccgctgcgaagagacacagttgcttgaatggtgtgtagtgcacaaga	3176
Db	3271	agagcgtcttatgacccctctgtcgtgttcagtgctaatgtatgtagaagaagatagaaca	3330
OY	3177	attgaatgtaagtggtcctccgttccccaataagcaatcgtcttccaaaggccctgtcctt	3236
Db	3331	cttgaatgtccagtggtcccgagacacccctggtgcacgtgtctcagaggaacccatccgtt	3390
OY	3237	caactgcagacattcgttaagacaatcgcgtcttggtagcaacagccgtgtgtgttcattaga	3296
Db	3391	tgcattgcagcttgcgcgaagacattgtcctatgtagagcaacagccgggtctgtacacatga	3450
OY	3297	tgaagatcaaaagacgcgaatctgcaggaatactcatcttatgtgaaggtctccgtga	3356
Db	3451	agagatatgtagggcagccgaaggagcaacatataccacttcatcgtagacatccctcgtga	3510
OY	3357	gaataacaacacacagtttggactgaaaggagcacagcttctctgcgcgcagaaacaaag	3416
Db	3511	gaataacaacacccagagtagagacaagaagaaacccagctctcctgtgtgcagaaacaaagc	3570
OY	3417	actgcattatgcaagggctctctcccaaaaaccaaatttatgtgtgtagaagcagac	3476
Db	3571	catgtccatagctcgcgtcctctgtttagacagccctcatatttgccttvtgtagaagactac	3630
OY	3477	ttcagccctcgtatatacagttgagaagagtggtgttcagcatcgtccctgtgataaagccagagac	3536
Db	3631	atcagctctgtgatatcagaaagtagaaaggtgtgcacaaagacccctgtgcagacaaagccagga	3690
OY	3537	gggaagagacatgctagttgtgtcacctcagagcgtcctctgcgaattcagaacgcagattgat	3596
Db	3691	aggcgcacactgtcattgtgtaccccccgcgctgtgcacacatccagaaatgcagatttat	3750
OY	3597	agtggtttctgacaatggaagaagataaaggagacaggaactcatcaagaagctctcttgaaga	3656
Db	3751	agttgtgttttcaagaatggcaagtccaagtgagcagctgtgcacacatcaacacagctgtcggccca	3810
OY	3657	tcgagacatatatttaagttatgtaagtgcacagttcag	3694





QY	1179	cgccttcgcttcgctccaaatgagcagtgaggaaagatacagtgagccagctctcgcagaagtt	1238
Db	1291	ggcgcttgcttgaggaaacagctgcgtgcggaaagacacagccgcgtgcagctatgcagaagct	1350
QY	1239	ataagatcccgatgagatgagctttatcaatctggtatgataatgacatcacagcttaaatgt	1298
Db	1351	ctaagacccccacagatgagcatgctcgtatctgatgacagagacattagaccataaattgt	1410
QY	1299	ggcgcatcatcagagacataatcttgaaatggttgatcaagaagccgtgtgttcgcggacac	1358
Db	1411	aagcgaccttcggggaaatactcgtgctggatcgaaagccctcggtgttctgcacac	1470
QY	1359	catcaatgaacaatatcaatgatgatgacgagatgatctgacataatgaaagatgagagagc	1418
Db	1471	gatacgctgaaataacatctgcgtatgcccgcgaaatgtaacacatgagatgagatgaaagc	1530
QY	1419	agcaagggaagcaaatgycgtatgattatcatctgagatcttccataaattbaaacat	1478
Db	1531	tgtaagaagagccaatgacatgacatatttatactgaaactaactaaatattgacact	1590
QY	1479	ggtatgggaaanaagggagcttcaaatgagtgagggagcagaagaacgagatcgcattgtccg	1538
Db	1591	ggtctgagagagaggggccagcttgagtggtgagaaacagaaatctgcattgtccgcg	1650
QY	1539	tgccctagcttcgaaaccccccaagattctgattttagatgagctatgcctgcgtgattc	1598
Db	1651	ggccctggttcgcgaaccccccaagatctctctgctgataagagaaatgtaactctcggaac	1710
QY	1599	agaaagcaagctcagctgttcaagctgacacttgagagaaagcagcaaaagctcgacatacat	1658
Db	1711	tgaaagtgaaagcagctggttcaagtggtcccttgataaagcgaagaaabaagcgcgactacat	1770
QY	1659	cgttgtagcagcacccgcactctctcaattctgaaggtgagattgattgtagaccctaaaga	1718
Db	1771	tggtatgagcccaatcgcttgctctacagttccgtaatgcgagtcgcatcttggttttgatga	1830
QY	1719	tggaatgctgcggagagaaagagcaaatctgtaacataatgagcaaaacgagctatata	1778
Db	1831	tgaggatcatcttgaggagaaagaaataatgataatgacatccatgaaaggaagagcatctact	1890
QY	1779	ctcaactgtgatagtctcacagagataataaaaagctgatga-acagatggagttcaatgacat	1837
Db	1891	caaacctgtcacaaatgcagacaaagagaaatgaaattggtattgaanaatgcacatgga	1950
QY	1838	attctactgaaagaagaagaccaactccctccctcgtcactctgtgaagag-----cat	1889
Db	1951	atccaaagtgaaagtgatgctcttggaataatgcttccaaagaattcaagggtccagttat	2010
QY	1890	caagctcagactcatgtaacagagcttgaggaatcccccacatcaagaagataagctt--	1947
Db	2011	aaaaagaagatcaactccgcagagatatatagtcacacacaagaagccagaagcagaagcttg	2070
QY	1948	-----cctgaagctctctcatataaatttaagtt	1979
Db	2071	tacaaagaggacttgatgataatgatactccatccagttccctctcgtgagatcttgaagct	2130
QY	1980	aaacaagcctgaaatgaccttgctggtcttcggggacatgctctctgttctaataatgaaac	2039
Db	2131	gaactcaactgaaatgaccttattcttggttatgatatatttgctatataaacaagaggg	2190
QY	2040	tggtcatcagagtatcttccataactcttgcaaaattaaataaccatggtt---ggaataaa	2096
Db	2191	ccctgcaaccgcatttcaataatatttccaagaagatlaaggaatcttaaccggagatga	2250
QY	2097	tgataaacaacataaagaatgatagtgcagaaattatitccatgatatctgtoaatitggg	2156
Db	2251	ggaatcctgaaacaaaagacagataagtaacatggtttctgattgtttctagctctgg	2310
QY	2157	tgatattgtcttgctcagttattcaatcgaaggatatttatcgcgacgaagcaggagaaat	2216
Db	2311	aattattcttattatacatcttctccagaggtcttcaaatitggaagcctggaggaaat	2370

QY	2217	lttcaagatgatgaatgaacacacttgcgcttcaagccaatggtatataatcagatatgtcgtg	2276
Db	2371	cctcaatgaacgsgcttcgatacatactgttctcagatccatctgcgtagaagcagagatgtcaagtc	2430
QY	2277	gtttgatgaaaaggaaaacgacgaagagctgtgacaacaatattatgacacatagataatgc	2336
Db	2431	gtttgatgaaacctaaaaaacacacccactggaagatgtgacaacaggcttgcgcaatgatgcgc	2490
QY	2337	acaaattcaagagcaacaaagcttccagatgtgcgtcttcaacaacaatgacaactaacat	2396
Db	2451	tcaagttaaaggsgcttatagttccaggtcgtgtgtcattaccacgaatatagacaactc	2550
QY	2397	gggaacttcaatcatcttcccttatataatggaatgggaatgacatcttcgatctgag	2456
Db	2551	tgggacacagcatattatatactccatctaatctaatgttgcgaattaaaccttttactcttagc	2610
QY	2457	tattgtccacgactctgcgctgtgacaggaatgatgtgaacccgacgaatgactcgtatgc	2516
Db	2611	aattgtaccctcatctgtcaataagcagaggtctgttgaataatgaaaatgtgtctcgtgacaagc	2670
QY	2517	caacaagaatgaagaagaactttagcatctgtgaaaagaatagacaactcgtgaagcttggagaa	2576
Db	2671	actgaaagaatgaagaagaagctagaagaagcctgtggaagatctgtctacagaagaaccatctgaaa	2730
QY	2577	tatacgttactaatagtgtcatttaaacaagggaanaagccttcgagcaaatgtatgaaagagat	2636
Db	2731	cttcggaactgtttgtcttcttgcctcgtgagcagaagtttgaatacatgtatgacagag	2790
QY	2637	gcttcaagactcaacaacgaanaatacctctgaaagaaagacacagatatattggaagctgttagc	2696
Db	2791	cttgcgaattacacatacagaanaactcttgaagaaagacacacatcttcggtgtctactttc	2850
QY	2697	attcagccatgcctttatataatttttgctatgagcagaggttttcgatttgagacattt	2756
Db	2851	tatcaaccagacgaatgtatatttttccatctgtcgtctgtttccggtttgtgcctactc	2910
QY	2757	aattcaagcttgcagcaatgacccccaagggcaltgtcatagatttctactgcgaattgtcata	2816
Db	2911	ggtgcgaatattgattccatgatacttccagatgtcttcttggtaattccagattatgctct	2970
QY	2817	tggagctatgtgcaltcgtgaaaaacgctcgttttgcgtcctcgtataatccaaagccaatc	2876
Db	2971	tgtgtcccatgtgcagttggtggagatgcagttcaatttgcctcgtcactatgcacaagccaagtc	3030
QY	2877	gggggctcgcggaatcgttttgccttgttggaaaagaaacaaatatagacagccagatca	2936
Db	3031	atcagcaagcccaacgtlcatgatcatgtatgaaaaagccctcgtatgtgaagatcaacagcc	3090
QY	2937	agaagggaaaagacgcagacaacatgttgaaggaaatttagagttctcgagaagctctcttct	2996
Db	3091	tcaagcgtctcaagccaataatcgtttgaaagaaatgtgacatttaatagatcgtgtctcaa	3150
QY	2997	ctatccatgtgccacagatgttctcatccctcgttgcctcgttgcctatccctcagatattgagcag	3056
Db	3151	ctatcccaactcgcacagacatctcccggtgcgtccacagggtcgtgaagctcctgaagtgaagaa	3210
QY	3057	aaagaacagtagcatttgttggggagcagcggcgtgttggaaaagcacttcgtttcaactctc	3116
Db	3211	ccagaacgtctgcctctgtatagtagtagcgtgtcgttgggaagaaacacagttgttcagctctc	3270
QY	3117	gaaagagactttatbaaccgcgtgcagaagacaaagctgtcttggatgtgtatgctgacaaa	3176
Db	3271	agaagcgtctcattatgaaccccttgcgttgcagtgacatattgattgacaagagataaaga	3330
QY	3177	attgaatgtatagctgtgctcgttccccaatatagcaatcgttctccaagacgtgtgtcctc	3236
Db	3331	cttgaatgtccaggtgctctcgagacacacccctgtggacatctgtctcagagacccactctgtt	3390
QY	3237	caactcgaagcatttcttgaagaaactcgcgtatgttggacaacagccgtgtgtgtccattaga	3296
Db	3391	tgaactgcagcattgtgcgaagacaatttgccttatgtgagaacaacagccgggtctcgtatccatgta	3450
QY	3297	tgaatcaaaaagccgcgaatgcagacaatatcatcttattatgaaagcttccctcga	3356

```

Db 3451 agagattgtagcagcagcaagagagccacatcacccacttcattcagacactccctga 3510
Oy 3357 gaatacaacacacaggtgactgaagagagacagcttcctgctgagcagacaagaag 3416
Db 3511 gaatacaacacacaggtgagagagacagacccagctctctggtgagcagacaagcgc 3510
Oy 3417 actagctatgcaaggctctctcccaaaaacccaaatttatgtgtgagagccac 3476
Db 3571 cattgcatagtcgcgcctctgttgacagcctcatatttctgttgatgaagctac 3630
Oy 3477 tttagccctcgtataatgacaggtgagaagtggttcagcctcgtatgaagccagac 3536
Db 3631 atcagctctgagatacagaagaagtgtaaaagtggtccagaagccctgagacaagccagaga 3690
Oy 3537 gggaagacatgctcctatgctgctacacagcctctcgtcaattcagaacccaagattgat 3596
Db 3691 aggcgcagctcgtatgtgtgtgcgccacgcgtgtgccacatccagaatbcagattaat 3750
Oy 3597 agtgtctctgcacaatggaagaataaagaacaagaacatccatcaagagctcctgagaa 3656
Db 3751 agtgtgtcttcagaaatgcaagtaagaagagagcatgacacatcaacagctgtgcgcca 3810
Oy 3657 tcgagacatatatttaagttaagtgaatgcacagtcag 3694
Db 3811 gaaagcatctatatttccatgctcagtgccaggtcgtg 3848

RESULT 4
AAD03488
ID AAD03488 standard; cDNA: 4279 BP.
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
DE Dog: P-glycoprotein; PGP; multidrug transporter; MDR1;
KM drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX
PN W0200123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000MO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
DR WPI; 2001-235373/24.
DR P-PSDB; AAE00303.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids; useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 3; Page 58-63; 111p; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are

```

```

CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC sequence is also referred as Genotype C cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Query Match 33.3%; Score 1232.8; DB 22; Length 4279;
Best Local Similarity 61.2%; Pred. No. 0; Mismatches 1257; Indels 132; Gaps 7;
Matches 2189; Conservative

Oy 162 caggtgaccctgtattatgttgataagtggtgctgctgtattttgttatacata 221
Db 358 caggtatgctattattattacagtggtgagcgtggtggtggtgctgtatcatacca 417
Oy 222 gatttccttggtatataactgcagacagacagacaagaagatcogaaaacagtttt 281
Db 418 gtttcattctgtgctcgtgcagcagagaagagagataactcaaatagaacaatttt 477
Oy 282 tcattcagtttgcacagagcagcagctggtttgatagcctgtacatcggtgaactaa 341
Db 478 tcattcattatgcagagagagatggtgtgttgacgtgcagtcagcttggtgagactaa 537
Oy 342 cactgcagcagacagactt---acaaatcagtgatgattgagaaatgacttct 398
Db 538 caccogctcacaagacagatgctctccaaatcaatgaagaattggtgcacaaattgcat 597
Oy 399 gtgtttcaaaaacatgtactatttcgatgtgctcgtcagtggtgttgtaagagctg 458
Db 598 gtctttcaactacatagaacattttccacggtttatagtggtttacacgtgtgtg 657
Oy 459 gaaactacccttagtgcattcatccacagtcctccttataatgcttaaggcgagcatg 518
Db 658 gaagtaacccttgatgtatttggccatcagccgttcttggaactttcagccgcatctg 717
Oy 519 ttctagagtgctcatcctatgcacagtaagaatlaagtcctattccaaagctggtgc 578
Db 718 ggcacaagatactatcttacttactgcataaagaactcttgccatgaaagctggtgac 777
Oy 579 tgtgcagaagaagctctgtcatcaatccgaacagtcataagccttaggtgcccagagaa 638
Db 778 agtagctgaagaagctcttagcacaatcagacgtgtgttgccttggtaggacaagaaga 837
Oy 639 agaactcaaaagctcttccctttaataatacaagaatgcttggtttattttcccca 698
Db 838 agaac----- 842
Oy 699 gtgctactaagtgtgttctgttnttgtlaagratcacagaatcacaagatgcaaa 758
Db 843 -----ttgaaggtacacaacaaaatttagaagaagcttaa 876
Oy 759 ggaatttgcaaaaagactatagcttcaaaagtgtcctctgtgtgtgtactctt 818
Db 877 aggaatttggttaagaagaagctacagcgaacacattctatgtgtgctgttcttatt 936
Oy 819 tatgaatgaactatggaactgtcttggtagtgaaccccttcattgaatgaaggaga 878
Db 937 gatcatgcatcatatgtctgtgcttctgtatggaagcctcctgtgctctccag--- 993
Oy 879 acctgataaccatcgagagctgtctgctgttcttcttgaatcaatagacagta 938
Db 994 ---tgaatatctatggaacaagaatcactacgtctcttcttctgtataatgggcttt 1050

```

QY	939	tgcgatttgagcagcagctcccttcacttcgtgaaaccttcgcaaaagcccgagagctgcctt	998
Db	1051	tagtatctgcagcagcactccccaagcattgaagcatttgaacgcgaagagcagcttia	1110
QY	999	tcattttccagtgctatttgataaagaaccocgataagatactttccacagcttgata	1058
Db	1111	tgaacttcctaaagataattgacaataaaccaagcattgcagctacttcgaagagtgagca	1170
QY	1059	taaacggtaaccataagaaagaaactgtggaattttaaanaatgtcttccttcatttcac	1118
Db	1171	taaacccgataataatctaaagggaatttggtaattcaaaaatgttcacttcagttaccctc	1230
QY	1119	aagacccatacaagattctcgaagaagctctgtaactcgaaattaaagtcctggagagacgt	1178
Db	1231	tcgaaaaagagtttaagattcaagctttaaagctctcaacctgtaagtgtagagtcagagct	1290
QY	1179	cgccctgtctgcgtctcaattgacgctggggaaggtacggtatgtccagctcttcgcaaggtt	1238
Db	1291	ggcgctgtgttgggaaacaagtgcgcgaggaagacacgacccgtgcagctgcatagcagaagct	1350
QY	1239	ataagatccggaatgataagctttcactgtgttgatataagaaatgcacaaagcatttaagtt	1298
Db	1351	ctatgaccccaacagatcggcagtgctgcgtacttgatcggacagcacaatgagacaataatgt	1410
QY	1299	ggcgcatlatacgagaccataattggaatgtgtatgcagaagccgtgtttgttcggagaccac	1358
Db	1411	aagcgactcttcgggaataattacactcgtgtgtgtgtgagccagagcctgtgtgttgcacacac	1470
QY	1359	catcatgtaacaataatacaagtatacgaatgacgagataatgtgactgtatggaagagatcggagagc	1418
Db	1471	gatagcctgaaaaaacttcgcataatgacgcgacgaataatgtcacacaatgagatctggagaagac	1530
QY	1419	agcaaaagggaagacaatcgcagatgatttttcctctgagatcttcctaataatttaacactt	1478
Db	1531	tgttaaggaagcccaatgcactatattttatcatgaataactcaactaaatttgcagacct	1530
QY	1479	ggtacgggaaagagagagctctcaaatgagatgagagcgcaaaacagaaatcgcgaattgtcgcg	1538
Db	1591	ggtctggagagagagggggcccaagctggagtggtgagcaaaacagagaaatcgcgaattcgcgcg	1650
QY	1539	tgccctatgttcgaaaccccagaatctctgattttagatgtagagcgtacgctgcgcctgattc	1598
Db	1651	ggccctgtctcgacaaccccaagaattctctctgcgtgtagggcaacgctcaagcctctgagacac	1710
QY	1599	agaaagacaagtgcgcgttcgttaagctctgcactctgtagaagagcgagacaagctcgcgactacaat	1658
Db	1711	tgaagaatgaaagcagtggttcaagtggtgcctctgtagaataagccagaaagccgagctaacat	1770
QY	1659	cgtctgtagcagcacgcgaacttctactatctctgaaagtgcagaaattgtagcttgcgaacctaaagga	1718
Db	1771	tgttagatgccatcgtgttgcgtactacagttctgttaactgcgaatgcattgtcgttlttgatga	1830
QY	1719	tgggaatgcctggcggagaaagagacacatgcctgcgaactaatgycaaaacagagctataatta	1778
Db	1831	tggagatcatctgtgagaaagaaatcatgatagcactcatgaaagagaaaggcgatttactt	1890
QY	1779	ctgcctgtgtagtctgcagagataataaaagctgata -acaagtgtgagtlcactaagaat	1837
Db	1891	caaatctgtccaactgtgaagacaagaggaataatgttagttgataaaatgtccactgtgtga	1950
QY	1838	attctacttgaagaagaaagccaactcactctccctctgcagctctgtgaagag-----cat	1889
Db	1951	atccaaaagtgaaagtgatgtcctcttgaaatgtctccaaaagatctcaaggtccagtttaat	2010
QY	1890	caagtcagactcaattgcacaagctgcaggaatccaccccaactaaagataagctct--	1947
Db	2011	aaaaagaagatacaactcgcagaggtatacatatgcacacaaagcgcaagcagaagaagcttgg	2070
QY	1948	-----cctgaagtctctcatataaanaatttaagtt	1979
Db	2071	tacaaaagagacttgatataagatgttactccgaagcttctctctctgtgagaaattctgaagct	2130

QY	1960	aacaacagcctgaaatgcgacctttgtgttctctggtggagacatgagctctctgtcttcaatgtgaac	2039
Db	2131	gaactccatgaaatgagcgcttattcttctgtgttggtatattctgtcatattcaataacaggag	2190
QY	2040	tgtctacacagatattttccatctcctcttgcaaaatataacacagttt---ggaaataa	2096
Db	2191	ccgtgcacacagcatcttccaataatatttccaagatataaggatccttaaccggagatga	2250
QY	2097	tgaataaaaccacatlaaagcatatgacagaatatttccatgatatctgtcatatttgg	2156
Db	2251	ggaatcccgaaacaaaacagacagaatagtaacatggttctctgtatgtttctacgtccctg	2310
QY	2157	tgtatttgccttgtcagttatttcaatgcagagatatttaacggcagacaggaggaat	2216
Db	2311	aattattcttttattcatcttttctctccagcgcttcacatcttgcaaaagccttgtagagat	2370
QY	2217	tttaagatgagatlaaagacacttgcccttcaagccatgatatataaggtatgtccgt	2276
Db	2371	ccctacataagcgcttcctgatacaatggtcttcacatccagctgcagacaggaatgtaacgtg	2430
QY	2277	gtttctgataaaagggaaaacacgcagcagaggctgtgacaacaatattatgcacataatagc	2336
Db	2431	gtttgattgacaccttaaaaaacacacacttgagacatgtaacacagcgctgcacatgatgcgcg	2490
QY	2337	acaaattccaaggagcaaacagtgctccagatgtgcgttcttaacacaaaatgcaactaacat	2386
Db	2491	tcaaagttaaaaggcgctatagttcttcacagcgcttgcgtgtctatccacagaaatatagcacatct	2550
QY	2397	ggagacttcaagttatctactcttcttcttataatgagatgtggagatgcagatccctgattcgcg	2456
Db	2551	tggagacaggaattatatacccttaacatacgtgtctgacataaacaacttaaccttaac	2610
QY	2457	tattgtcccaagtaactgcgtgcgaagaatgattgtaaaacgcagcaacatgacatgactgtgc	2516
Db	2611	aattgtaccacatacttgcaatagcagagtggtgtgtgaatgaaatggtgtctgcgaacaagc	2670
QY	2517	caacaagaataagcaagaacttaagcatgcgtcgaaagatagcaactgaaagcttvggagaa	2576
Db	2671	actgaagaataagaagaagcttagaagagcgttggaaagattgtctacagaaagcatccatga	2730
QY	2577	tatcgtctactatagtgctcattaaacagggaacagccttcgcgcgaatgtaatgtaagagat	2636
Db	2731	cttccgcgaactgtgttcttcttcttgacgcggagacaggtttgatacatgattgcagagag	2790
QY	2637	gcttcagacatacaacagaatactccctgcgaagaagacacagattatctgaagactgttgcgc	2696
Db	2791	tttgcgaagtacacatacagaaacctcttgaggaaagacacacatccctcggggtccctatttc	2850
QY	2697	atcagccatgctcttatataattttgcctatgcctagcagaggttttcgatttggagcctattt	2756
Db	2851	tatccccaagccaatgattgtattttcttcctatcttcctatctgcgttccggttgcgcctactt	2910
QY	2757	aattcaagcttgcgaagatgtgccccagagggcagatgtctcatagtttttactgtcaattgcata	2816
Db	2911	ggctggcaaaatgagtctcatgtaacttccaagaaatgtctcttggtaattccaagcatgtctt	2970
QY	2817	tggagactatggccaatcggaanaaagcgtctgtgtgtctctgtaataattccaagccaatc	2876
Db	2971	tgtgtgcataitgcagtggggcaggtctgaatttcaattgtctcctgcctatgcccacaaagccaagt	3030
QY	2877	gggggcttgcgcatctgttttgccttgttgcgaanaagaaaccaaatalagacagccgcatga	2936
Db	3031	atcagcagccacagctcatcatgtatcatctgaaanaagccctctgcattgcagctatagacc	3090
QY	2937	agaagggaaaaagccagaaacatgtgaaaggaaattcagagtttctgagaagctctcttctt	2996
Db	3091	tcaagcgctccaagcaaaatcgtttgaaaggaatgtgcacatttaatgaggtctgttctaa	3150
QY	2997	ctatccatgtctcccaaatgttttctacctcccgcttgcctatcccccagtatgtgaagcgag	3056
Db	3151	ctatcccaacttcgacacagacatcccgcgtgtctccaggggctctgagccctcgaggttgaagaaagg	3210
QY	3057	aagacagttagcatcttcttgggagcagcggtcttgggaanaagacacttctgttcaactctt	3116

[illegible]

XX Stocker PJ, Steinel-Crespi DR, Crespi CL, Reif TC, Patten CJ;  
 XX  
 XI WPI; 2001-235373/24.  
 DR  
 DR P-PDSB; AAE00304.  
 XI  
 PI New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PI for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX  
 XX Claim 1; Page 66-72; 11pp; English.

---

The invention relates to dog P-glycoprotein (PGP) also referred as multidrug transporter (MDR1), and nucleic acids encoding them. The invention also includes fragments and biologically functional CC variants of dog P-glycoprotein. PGP and their nucleic acids are CC useful for determining the bioavailability of drugs and for CC screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterised by PGP activity, by CC reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.

The present sequence is dog P-glycoprotein (PGP) cDNA. The CC PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.

SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

```

Query Match      33.2%; Score 1229.6; DB 22; Length 4317;
Best Local Similarity 61.1%; Pred. NO. 0;
Matches 2187; Conservative 0; Mismatches 1259; Indels 132; Gaps

```

QY	162	caggttgaccctgatatgtcggatagtgttgcctcgatttttgttataca	221
DG			
DG	408	caagtagtcctattataacagtyggatctgctgcgctgcyggccgtcatcca	467
QY	222	gatctctgtgtatataactgcagcacgaacagaaggattcgaaaacagtttt	281
DG			
DG	468	gyttcatctcgtgctgcgacgaagaacagatactcaaattagaaaaaatltt	527
QY	282	tacttaagtttggcccagaacctgcgcgtgtttgttagctgtgaaatgcgtgaattaa	341
DG			
DG	528	tcatgtatcatactgcagagagatggcgtgttcgtgcagtcaagaagcttgggaacttaa	587
QY	342	catcgcatacagacagatcyy---acaaaacagtgtatgtlatgtgagaataattgctct	398
DG			
DG	588	caeccgycctcacagaagaagatgcccaaataatgaagaattggcgacaaaattggaat	647
QY	399	gtttttcaaaaacatgctaactlcttcgattggcctgcgtgcagttgtttggtgaaggctg	458
DG			
DG	648	gtttcttcaatcaatagaacaattttccocggttttatagtgggtttaaacctgttg	707
QY	459	gaaactcaccctagtgaactctaccaacgctcctcttatataitgctgcggcgacagcaty	518
DG			
DG	708	gaagctaacacctgtgattttgtgcacatcagcctgttcttggacatttcagccgcatctg	767
QY	519	tttaagatgtgcatactcatcattgacacgaagaaagattaaagtgcctataccaagctgggpc	578
DG			
DG	768	ggcaaaatacctctctcaattctacgtataaagaaccttggcctatacgaanaagctggagc	827
QY	579	tgtgcagaaagatctctgtcalcaatccgaaagaftacatgcctttaaggccccagagaga	638
DG			
DG	828	agtagctgaagaagctcttagcagcaatcagaactgtagtgcctttggaagacaaaagaa	887
QY	639	agaacttcaaaagtlcttccctttaataatacaagaatatgcttggtttattllccoca	698
DG			
DG	888	agaac-----	892

OY 699 gtggtactaagtgtgtctgttnttltgtaagtatcacagaatctcaaatgacaa 758  
Db 893 -----tggaaaggtcacacaaaatlttgaagaagctaa 926  
OY 799 ggaatttggcataaaagagactatagctcaaaagtctcttggctgtgacttctt 818  
Db 927 aagaatttggataaagaagactatcagcgccaacttcttattgttgcgcttcttatt 986  
OY 819 tatgataagaaactatgtaactgtcttlttgtatgtaaaactctctgtacttctaagaa 878  
Db 967 gatctatgatatatgtctgtgcttcttctgtatgtaagactcttctgtctctccag --- 1043  
OY 879 acctgatatataccatcggaagctgtctgttcttcttctgtatgtaatccatagacgta 938  
Db 1044 ---tgaataactatattgtaacaggtactcactgtcttcttcttctgtatatttgggctt 1100  
OY 939 ttgcatltggaagcagcagtcctcacttltgaaactctgcaatagcccgagagcttgcct 998  
Db 1101 tagtatgtaagcagcattccccaagcatltgaagcatltgcaagcagagagcagctta 1160  
OY 999 tcatatttccaggtatttgaataaagaccagtatagtaactcttccacagcttgata 1058  
Db 1161 tgaactctcaagataatgtgacataaaccagaagcatltgacagctaltcgaaagtggaca 1220  
OY 1059 taacctgaatccatagagaagcagctgtggaatttaaaatgttcttccatattccac 1118  
Db 1221 taaccagataataatgaagggaatlttgaaattcaaaaatgttcaacttcaagttaaccttc 1280  
OY 1119 aagaccatctatcaagatlttgaaaggtctgtaactcagaattaaagltctgagagacagt 1178  
Db 1281 tcaaaaagaagtaagtaagtaacttaagggtctcaacctgaaggtctcagaagcagcagct 1340  
OY 1179 cgcttgttgcgttccatagcagctgtggagaagtaacgtatgtccagctctcagaggtt 1238  
Db 1341 ggcgtgtgttggaaagcaggtgctgcggagaagcagcagcagctgtaagagagct 1400  
OY 1239 aatgtatccgagatgtggttcatcatagtgatgtaagaaatgacatcaagagctttaatt 1298  
Db 1401 ctatgatacccaagatgtgcatgtctgtatgtatgtgacagagacatlaagacataaagt 1460  
OY 1299 gcgcgattatcgagacatatttggatgtgtaagcagaagcctgttlttgcggagccac 1358  
Db 1461 aagcatctctcggaagaaatctactcgtgtgtgtgtaagtcagagcctgtgttgcacccac 1520  
OY 1359 catcagtaacaatatcaagatgtaagcagaaatgattgtactgtatgaaagatgtgagagac 1418  
Db 1521 gatacgtgaaacacatcgtctatgycgcgcaaaatgtcaccatgtatgtatgtgaagaag 1580  
OY 1419 aagcaaggagcaaatgctgtatgattttatcatgtgaggttccctaataatttaactatt 1478  
Db 1581 tgttaagaagcacaatgtcctatgtattttatcatgtgaactaactaataatttgcacactct 1640  
OY 1479 ggttaagggaanaaagagctcaaatgtagtgagggcagaagacagagatcgcaatttctcg 1538  
Db 1641 ggttggagagagagggccgcgctgtagtggtgagcagaagaaacagagatcgcacttgcgt 1700  
OY 1539 tgccttagttcgaaacccaagatctgattttagatgagcgtacgctgcgcctgtatcc 1598  
Db 1701 ggccttgcgttcgcaaccccaagatcttcttctgtgtagtgagcaacgtcagctctgacac 1760  
OY 1599 aagaagaagctcagctgttcaagctgtgactgtgagaagggagcaaggtccgagactacat 1658  
Db 1761 tgaagaagtgagaagcagtggttctcaagtgccctgtgataagggccagaagggccgagactacat 1820  
OY 1659 cgttgaagacacagacttctactactatcgaaagtgcagattgtatgttgaacccaataag 1718  
Db 1821 tgtgaagactcagtttcttctgtcagagcttgcataatgcccagatgtcagctgttcttga 1880  
OY 1719 tggaaatgtcggcgagaaagagcacatgctgtaactaatgtgcgaagaggtctatatta 1778  
Db 1881 tggatgtcatctgtgagaagaaatcatgtatgaactcatgaaagagagagcattactt 1940  
OY 1779 ttcaactgtgagtccacagatatataaaagctgtatga -acagatggaggtcaatgaat 1837

Db 1941 caaactgtcacaaatgcagacagaaggaatggaattggaattgtagaanaatgccactgtga 2000  
OY 1838 attctacttgaagaagaagaccacactcacttccctgtcactctgtgaag ---cat 1889  
Db 2001 atccaaagtgtgaagtgtgcttggtaaaatgtctccaaaagatctcagggttccagttaat 2060  
OY 1890 caagtacagcttcatgtgaagaagcttgaggaatctcaccacaatctcaagagata ----- 1941  
Db 2061 aaaaagaagatcaaccgcagagagtatcacatgcacaaagccagaagacagaagcttgg 2120  
OY 1942 -----agcttccotgaagctcctcctatlaaaatlttaagtt 1979  
Db 2121 tacaagaagagacttgaatgagagatgttctcactgaatttcttcttcttgaggtcttgaagct 2180  
OY 1980 aacaagcctgaatggccttgtgtgtcttggagacattggtctcgttctaagtgaac 2039  
Db 2181 gaactcaactgaaatggccttatttgtgtgtatatttggctatatttgcattataaagcgaag 2240  
OY 2040 tgttcaacagtaatttccatcacttcttgcanaaatattacacatgtt ---ggaataa 2096  
Db 2241 ccgtgcacacagcatlttcaataatatttcaaggaattataggaatcttaccgagatga 2300  
OY 2097 tgaataaacacacattaaagatgatgcagaatatttccatgatatctgtcatlttgg 2156  
Db 2301 ggaactcgtgaacaaacagacagatagtaacatggttctcgtatgttcttctagcttgg 2360  
OY 2157 tgttattgtctgtcagatatttcatcagcaggaattatttaccgcagagcagggaat 2216  
Db 2361 aattattcttattacatttcttctccagggctccacattgtgcanaagcttggagat 2420  
OY 2217 tttaacgataagatlaagacacttggccttcaagcagatgttatcaagatattgtctg 2276  
Db 2421 ccccaactaagcggcttcgatacagtggttccagatccatgtgagacagagatgcagctg 2480  
OY 2277 gtttga tgaagaaggaacacagcagcagagagcttgacacacatatlttgccaatagataag 2336  
Db 2481 gtttga tgcaccccaaaaacacacacacagcttgagatgtgacacacagcttgcacatgtgcgc 2540  
OY 2337 acaaatlcaaggagcaacaggtccagatgtgcgttcaacacaaatgtcaactaacat 2396  
Db 2541 tcaagttaaagggtctataggttccagcgttgcgttcatatccagaatataagcaaatc 2600  
OY 2397 gggacttcaagtatatttcttcttataatgataatgagatgacatctcgttctggag 2456  
Db 2601 tgggacaggaatattatataatcccttaactatgtgttgcgaatcaactaacttactcctgaac 2660  
OY 2457 tatgtccagtaacttgcgtgacaggaatgtgaacccgcagcagatgactgatttgc 2516  
Db 2661 aattgtacccaatcatatgtcaatagcagaggttgttgaataatgaatgttgccttgacaagc 2720  
OY 2517 caacaaagataagcaagaacttaagcatgtctgaaagataagcaactgaaagcttggagaa 2576  
Db 2721 actgaagaagtaagaagagctagaagagcgtgggaagattgtctacagaacatctcgaa 2780  
OY 2577 tatagtactatagtgatcaatacaagggaagaaagccttcgacaaatgtatgaagat 2636  
Db 2781 ctcccgaaactgtgttcttcttgcctgcggagcagaagtttgaataatcatgtatgcagag 2640  
OY 2637 gcttcagactcaacagaaataatccctcgaaagagcagcaactatttggaaagcgttga 2696  
Db 2841 ttgcaagatccatacagaacactcttggagaaagcacacatcttcgggttccatcttc 2800  
OY 2697 atcagacatgccttataatatttgcctatgtaagcaggttccagatttggagactaatt 2756  
Db 2901 tatcaccagagcaatgatgtatatttccatagtctgcgttccogtlttgggttgccttga 2960  
OY 2757 aattcaagctggaagatgaccccaagaggtcagttcatatgttcttaactgtcatatga 2816  
Db 2961 ggtggaanaatgagttcatgaacttccagagatgttcttgttatctccagcatgttct 3020  
OY 2817 tggagctatggcatcggaaanaagcgtcgttgttgcctccggaatattccaaagccaatc 2876

```
Db 3021 tggctgcattgagcagtgaggcaggttcagttcatttctctctgactatgcccagaagc 3080
Qy 2877 gggggcctgcgcacatctgttctgttggaaaagaacaaatagacagccgacgtca 2936
Db 3081 atcagcagccacgcgtcatgatacttgaaaagaacccctctgtatctgacatctcaagccc 3140
Qy 2937 agaaaggaaaagacacagacatgttgaaaggaaatttagatctgcgaagaagctcttctt 2996
Db 3141 tcacgcgctcaagcacaataacgttggaaaggaaatgtgacatttaagtatgagtcgcttcaa 3200
Qy 2997 ctatccatgtcgcagcagaatgttttcatctccctgctgtgcttaccctcagttatggcagg 3056
Db 3201 ctatccacacgcagacacatccccgtgtctccaggggctcagcctcagggaggaagg 3260
Qy 3057 aaagacagtgacttlttgaggagcagcgctgtgggaaaagcactctcttcaattct 3116
Db 3261 ccagacgctgtgcccctcgttaagcagatgctgtgggaaggacagattgttccagctcc 3320
Qy 3117 gcagagactttagaccctgtgcaagacaaagtctgttcttgatgtgtgatacgcaaaaga 3176
Db 3321 agagcgtctctatgacccctgtgctgttccagtgtaattatgtgcaaaagataaagca 3380
Qy 3177 atggaatgtacagtgtgctcgttcccaataagcaatcgcttcccaagaagcctgtgtctt 3236
Db 3381 cctgaatgtcagtggtgtcgcagacacacctgtggtcgtctcagagccatctgtt 3440
Qy 3237 caactgagcagctgtgcaaacatcgctatggtgtgacaaagcgtgtgtggtgcattaga 3296
Db 3441 tgaactgacgactgtcggaacatgtgctatggaagacacagcgggtgcgtatcatacaga 3500
Qy 3297 tgaagataaagaagccgcaaaatgacgaacaaatccattctttatggaaagcttccctga 3356
Db 3501 agagattatgcaaggcagcagcaagaggaacacacacacttccgagacatccctga 3560
Qy 3357 gaataaacaacacaaagtgtgacgtgaagagcagcagcttctgtgcccgcgaagaaacag 3416
Db 3561 gaataaacaacacagagtagaagaaagaaagcagctctctgtgtgcagaaacagcg 3620
Qy 3417 actgacatgtcaaggcgtctctcccaaaaacccaatttatgtgtgataggccac 3476
Db 3621 catgcatagctcgtcgtctgtgttagacagccctcattatgttcttggatgaagctac 3680
Qy 3477 ttccagccctcgataatgacagtgagaagtggttccagcatgccccttgataaagcagagac 3536
Db 3681 atcagctcgtgatacaagaagtgtgaagaaagttgtccaagaagccctgcgaagcagagga 3740
Qy 3537 ggggaagacatagcctatgtgtgctcaactcaagcgtctcttcaattcaagcagcagattgat 3596
Db 3741 aggcgcgaacctgtcatgtgatacgccacgcctgtgtccacatccagaatgtgagattat 3800
Qy 3597 agtgttctgcacacatggaagaataaagaaagaaagaaactcaagagccctctgagaaa 3656
Db 3801 agtgtgttcaagaatgtgcaaaagtcgaagagacatggcacaacacagcgtcgtgtccca 3860
Qy 3657 tcgagacataatatttaagttagtgaatgacagtcag 3694
Db 3861 gaaagcatctattttccatgcatcagtgctcagagctg 3898
```

```
RESULT 6
AAH57442
ID AAH57442 standard; cDNA: 4349 BP.
```

```
XX AC AAH57442;
```

```
XX DT 10-SEP-2001 (first entry)
```

```
XX XX Human intestine cell specific cDNA sequence SEQ ID NO:282.
```

```
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
```

```
XX OS Homo sapiens.
XX PN W0200132927-A2.
XX PD 10-MAY-2001.
XX PF 02-NOV-2000; 2000WO-US30396.
XX PR 04-NOV-1999; 99US-0163508.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Sornasse T, Sellhammer JJ, Watson GA;
XX DR WPI; 2001-291057/30.
XX PT New cell and tissue specific polynucleotides useful for diagnosis,
XX PT prognosis or monitoring of treatments for disorders where the gene is
XX PT associated with a cancer, immunopathology or neuropathology -
XX PS Claim 1; Page 207-208; 327pp; English.
XX CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
XX CC sequences (I). (I) can have cytostatic, immunomodulatory and
XX CC neuroprotective activities, and can be used in gene therapy. (II) and
XX CC proteins (II) encoded by them are used in high throughput screening
XX CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
XX CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
XX CC their fragments, immunoglobulins, inhibitors, drug compounds and
XX CC pharmaceutical agents. Expression of (I) in a sample indicates the
XX CC differentiation of embryonic stem cells into a tissue selected from
XX CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
XX CC tissues. (I) and (II) are used to produce an expression profile that
XX CC defines a metabolic or developmental process, treatment, condition,
XX CC disease or disorder. The gene profile can be used for diagnosis,
XX CC prognosis or monitoring of treatments and for investigating a
XX CC predisposition to a disorder where the gene is associated with a
XX CC cancer, immunopathology or neuropathology.
XX SQ Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Query Match 33.1%; Score 1225.2; DB 22; Length 4349;
Best Local Similarity 60.9%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 1268; Indels 132; Gaps 6;

Qy 162 caggttgacctgtattatgttggaatagtggtgtgctgcttgatttggttacataca 221
Db 461 caggtatgctctattatcagtggaatcgtgtggtggtgctgtgtgtgtacatcca 520
Qy 222 gatttctgttgattataactgacgcagcagacgaagagattcgaaaacatttt 281
Db 521 gtttcaatttgggtcctgcgcagcgtggaagacaaataacaaaattgaagaaacatttt 580
Qy 282 tcatcagtttggcagacagacatcgctgtgttgatagcgtgtgacatcgtgtgaactaa 341
Db 581 tcatgtataatgagcagagagatagcgtgtgttgatgtgcagcagtggtggagactaa 640
Qy 342 cactcgcataagacatgac---aaatcagtgatgtgatgtgaatgaatgacct 398
Db 641 caccgacttaagatgagtgtccttaagaattaaagattatgttgcaaaatggaaat 700
Qy 399 gttgttcaaaacatgcttaacttttcgattgacctgcctgcagtggttggtaaggcgtg 458
Db 701 gttcttcaagtaatgcaacaacttttcaacgtggtttatagtaagattacacgtgtgtg 760
Qy 459 gaaactcacactagtagactatccacgctctccttcaataatggtctcagcggcagcag 518
Db 761 gaagctaacaccttgatttggccatcagctcgttcttggacgtcagcgtcgtcgtg 820
Qy 519 ttctagagtgatcatcctaatgtacagtaagaagaaatgaagtgcctattccaaagctggggc 578
```



Db 821 ggcgaagatcattcattactactgataaagaaccttagcgta tgcgaaagctggaac 880  
Qy 579 tctggcaagaagctctgtcattcaatccgaacagcttagccttagggcccgaga 638  
Db 881 agtgcggaagagctcttgcgcagcaatcagaactgagatctgagagacaaga 940  
Qy 639 agaatctcaagagcttctctttaataataacaagatgcttggtttatttccca 698  
Db 941 agaac----- 945  
Qy 699 gtgcctactaagctgtgtctgttnttctgtaaggtatacacagaatccaaagtcaaa 758  
Db 946 -----ctgaaaggtacacaaaatttagaagaactaa 979  
Qy 759 ggaatttgcataaaagagactatagcttcaaagtctctctgtgtgtgtactctt 818  
Db 980 aagaatttggataaagaagactatcacgccaatttcttaagtgctgcttctctgt 1039  
Qy 819 tatgaatgaaacctatgagactgtcttctgtatggaacctcctgattcttaacgaga 878  
Db 1040 gatcattgacattatgctctgcgcctctgtgtatggaacctgtgctctcagggga 1099  
Qy 879 acctggaataccatcggagactgtctctgtgttcttcttaagttaatccatagacgta 938  
Db 1100 -----atattctatgacacaagctactcacgttacttcttcttatttaattgggcttt 1153  
Qy 939 ttgcatgtggaacgagcctcctcacttctgaaaccttcgcaaacgcccagggagctt 998  
Db 1154 tagtgttggaaagcatcccaagcatctgaaagcatctgcaaaagcagggagcgctta 1213  
Qy 999 tcatatttccaggttatctgtaaaagccagatagataaacttctccagcgcgagata 1058  
Db 1214 tgaactctcagaataatctgataataaagccaagtattgacagctattcgaaagctggga 1273  
Qy 1059 taacctgtatccatagagaagaaactgtggaatttaaaaatgttcttccatattccatc 1118  
Db 1274 caaacagataataatttaagggaatttgaattcagaatgttccactcagttaccatc 1333  
Qy 1119 aagaccattatcnaagatctcgaagaagctcgaatctcgaatataagctcggagaaacgt 1178  
Db 1334 tcgaaaagaagltgaagatcttgaaggccttgaaacctggaagctgagcagaaacgt 1393  
Qy 1179 cgccctgtcgtctcgaatgagcagtggaagagatagcttagccagctctcgaaggt 1238  
Db 1394 ggcctgtgttggaaacagtgctgtgggaagagcacaacagctccagctgagtcagaagct 1453  
Qy 1239 atatgatacggatgatagtgcttatcatatggttgagatgagaatgacacagagctttaagt 1298  
Db 1454 ctatgaccccaaggggagtgatgtgtcagtgatgagacagatatlaagacataaagt 1513  
Qy 1299 ggcgcatattcgaacccatattgagtggttcaatcagaagcctgttcttgcggagcac 1358  
Db 1514 aaggttcttaaggggaatcattgtgtgtgagtcaggaacctgtattgttgcgacccac 1573  
Qy 1359 catcgataacatcgaatctgagcagagatgattgactgtagaagatgagagagc 1418  
Db 1574 gatgcttgaaaaaacttgcctatgctgcggaagatgtcaccaatgagatgagaagagc 1633  
Qy 1419 agcaaggaagcaaatgctgataattatcatatggaagttccctaataattatcatt 1478  
Db 1634 tgtcaaggaagccaatgcttacttatactatgaaactgcctcataaatttgcacacct 1693  
Qy 1479 ggttgagggaagaaagagctcaaatgagtgaggggcagaagacgaagatctgcaattgtcg 1538  
Db 1694 ggttggaaagaggggcccaggttgaatggttggcagaagcagaagatctgcacattgca 1753  
Qy 1539 tgccttaattgaaacccaagatctgatttgaatgagagctgctgcgcctgattc 1598  
Db 1754 tgcctgttctgcaaccccaagatctcctctgtgataggaagcgaagctgcagccttggagac 1813  
Qy 1599 agaaagcaagctcagctgttcaagctgacatggaagagcgagcaaaagctcgaactacaat 1658  
Db 1814 agaaagcgaagcagtggtctcagtgctgtgataagcgcaagaaagctcggacacacat 1873

Qy 1659 cgtgtgagcacaccgacttctactactcgaagtgcagatttgattgtgacccaaagga 1718  
Db 1874 tctgtagctcatcgttcttctacaggttcgaatgctgacgctgctgctgttctgata 1933  
Qy 1719 tggaaatctgcggggaagagagacatgctggaactatgycgcaaaagagctatata 1778  
Db 1934 tggagctattgttggagaaggaatcatgatactgacatcgaaagagaagacattact 1993  
Qy 1779 ttcacttgtatgtcagagatataaaaaagctgtatgacagatgagatgacata 1838  
Db 1994 caaactgtcacaaatgacagcagcaggaatgaagtgtgaattgaataatgacgttgata 2053  
Qy 1839 tctta-----ctgaaagaagacacacactcactcctcctgcagctgtg 1881  
Db 2054 atccaaagatgaattgatagtccttggaaatgttcccttaaatgtatcagaatccagctcaat 2113  
Qy 1882 aagagcatca-----agtcagacttcatgacaagcgtgagga 1919  
Db 2114 aagaaaagaatcaactcgtgaggaagtccgttgatcacaagcccaagacaagaagcttag 2173  
Qy 1920 atccaccaatctaaagataagctctccctgaagctctctctataaaaaatttaagct 1979  
Db 2174 taccaaagagctctgagtgaaagatacctccagcttctccttggaggagattgaaagct 2233  
Qy 1980 aacaagcctgagtgcccttgggttgggttgggagacatgagctgtcgtcttaattgaaac 2039  
Db 2234 aaatttaactgaaatggtccttatttgtgtgtgtgtatttggcattatataaaygaa 2293  
Qy 2040 tgtcacagataatttccatcatccttctgcaaaatataacatgtt---ggaataaa 2096  
Db 2294 cctgcaaacagcatcttgcaataatatttccaagaatataaggttttacaagaattga 2353  
Qy 2097 tgataaacacacataaagcatgagtcagaagaatttatccatgatatctgcatttggg 2156  
Db 2354 tgactcgtgaacaaacagcagagatagtaactgttcttccactatgttcttctagcccttg 2413  
Qy 2157 tgttattgcttgttcaagtatttcatgaggaattatttaaggcagagcgaaggaaat 2216  
Db 2414 aattattcttatttatacatcttctcctcagaggttccaacttgcgaagcgtggaagat 2473  
Qy 2217 tttaacgataagataaagacacttgccttcaagaacatgtataacagagatatctctg 2276  
Db 2474 cctcacaagcggctcccgataaagtttctcgaatccatgctcagaagatgtgttg 2533  
Qy 2277 gtttgaagaaaggaacacagcagagagcgttgacacaataattgccaatagatagc 2336  
Db 2534 gtttgaagcccttaaaaaacacacactgagcatgtactacagcgtcgcgaatgagtgc 2593  
Qy 2337 acaaatccaaggaacacaggttccaggttgcgctcttaaaccaaatgacataacat 2396  
Db 2594 tcaagtaagaagggtcatalaggtccaggttgcgtctaatttaaccaaatatagcaatc 2653  
Qy 2397 gggacttcatgattcattccttctataatgataatgagagatgacatctcgtattgag 2456  
Db 2654 tgggaaggaataataataatccttccatcctatggttggcacaatacactgttactctagc 2713  
Qy 2457 tattgtccagtaactgccgtgacgaagaaatgataaaccgcgaacatgactgagattgc 2516  
Db 2714 aattgtaccatcattcgaatagcagaggttgttgaataatgttgtcttgagaagc 2773  
Qy 2517 caaagaagtaagcagaacttaagaatgcatgtcggaaagagatagacatggaagcttggaga 2576  
Db 2774 actgaaagataaagaagaacttagaagtgctggaagatctgcactggaagcaatagaana 2833  
Qy 2577 tatagctactatagcttcatlaaagaagggaanaagccttcogaaatgatatgaagat 2636  
Db 2834 cttccgaacagctgttcttcttgcacagagcgaaggttgaacatattgatatgctgaag 2893  
Qy 2637 gcttcagactcaacagagaatacctcogaaagaagcagagattattggaagctgtttagc 2696  
Db 2894 ttgcagtaaccatacagaacactcttggagaaagcacaacatcttggaaattacatttc 2953



```

OY 2697 attccgcaatgcttataatttgcctatgcagcagggttgcatttgagcctatt 2756
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2954 cttcaccaggaatgatattttctctatgctggttcggttttgagcctact 3013
OY 2757 aattcaagcttgacgaatgaaccccaaggcatttcattatgttttactgcat 2816
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3014 ggttgacataaactatgaacttgagagatgttcctgattatcttcagctgt 3073
OY 2817 tggagctatgcccacatcggaanaacgctcgttttgctcgtgaatattccaa 2876
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3074 tgggtccatggcctgtggggaagtcattcattgtctcctgactatgcca 3133
OY 2877 ggggctgcgcacatcgtttgcctgttggaaagaaaccaataatagacgc 2936
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3134 atcagagcccatcatcatatcatgtatgaaaaaaccccttgattgacagc 3193
OY 2937 agaaaggaanaaacacacacatgtgaaggaatttagaatttcogaaatc 2996
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3194 ggaagacctaatgccaacacatlgaaagaaatgtcacatttggatgaat 3253
OY 2997 ctatcatgtgcgcacagatgtttcattcctcgttgcttaaccctcagta 3056
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3254 ctatccaccccgacccgacatcccaatgcttcacaggagactgagctgt 3313
OY 3057 aaagacagtagcaattgtgggagcagcgtgttggaanaacacatcgtt 3116
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3314 ccagacgctgctcgtgtgggacagctgtgtgggaagacagctgtccagc 3373
OY 3117 gcagagactttagaccccgctgcagacacagtgctgttgatggttgat 3176
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3374 ggaagcgttctacgaccccttgcaagggaatgtctgtatggaanaac 3433
OY 3177 attgaatgacagtgctcgttcctccaaatagcaatcgttccaaagcgt 3226
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3434 actgaatgctcagtggtctccgagacacccgtgacatcgttccaggaag 3493
OY 3237 caactgacagctgtgcagacatcgcctatgtgtgacacacagcgtgt 3286
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3494 tgactgacagctgtgcagacatcgttgcctatggaacacacagcgtgt 3353
OY 3297 tggagtaagaagccgcaaatgacgaataatcattctttagaagttccct 3356
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3554 agaatatcgtgagggcagcaaggaagccacacataatgcttccatgact 3613
OY 3357 gaataacacacacaaagtgtgacgaaggaagcagcatttctgctgcga 3416
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3614 taatatagcaactaaatagtaggaacaaagacacagctcgtgtgcca 3673
OY 3417 actagctatgcaaggcctctctccaaaacccaatttattgtgtgag 3476
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3674 cattgcatagtcgtgctcctgttagacagcctcatatttgccttgg 3733
OY 3477 ttgcgctcgtatataatgacgttggaaggtgtgttcagatgccttgg 3536
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3734 gtcagctctggtatcagaagatgaaaggtgtcccaagaagccctgca 3793
OY 3537 ggaagacagcagctatgtgtcactcacaagcctctcgaattcagaag 3596
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3794 aggcgcgcacctgcatgtgtatgtctacacgcctgtccacatccaga 3853
OY 3597 agtgttctgcacaaatggaagaaatgaagaaacgaatcatcaagagct 3656
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3854 agtgtgttgcagatgacagatgaaggaagacatgacagcagcgtgtgc 3913
OY 3657 tcgagacatatatttaagtgaatgacagcagc 3694
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3914 gaaagcactcatatttccaatgctgtccagcgt 3951

```

RESULT 7  
AAO72872  
ID AAO72872 standard; cDNA; 4646 BP.  
XX

```

AC AAO72872:
XX 21-JUN-1995 (first entry)
XX Human multidrug resistance gene (MDR-1) encoding a P-glycoprotein.
DE
XX
XX Human; tyrosinase; compound; hair; follicle; liposome; lipophilic;
KW lipophilic; stratum corneum; melanin; hair dye; glycoprotein; alopecia;
KW chemotherapeutic; transforming growth factor; growth stimulant; aromatase;
KW cyclosporin A; macromolecule; polymer; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 425..4267
FT CDS /tag= a
FT /product= P-glycoprotein
XX
XX W09422468-A.
XX
XX 13-OCT-1994.
XX
XX 01-APR-1994; 94WO-US03634.
XX
XX 02-APR-1993; 93US-0041553.
XX 13-JAN-1994; 94US-0181471.
XX
XX (ANTI-) ANTICANCER INC.
XX
XX L1 L, Lshko VK;
XX
XX WPI; 1994-332816/41.
XX P-PSDB; AAO72872.
XX
XX Liposomes for delivering protein, nucleic acid etc. to hair
XX follicles - e.g. to restore hair colour, prevent hair loss during
XX chemotherapy, stimulate hair growth etc.
XX
XX Claim 8; Page 70-77; 100pp: English.
XX
XX The nucleotide sequence of the human multidrug resistance (MDR-1) gene
XX which encodes a P-glycoprotein. This is an example of a compound which
XX can be delivered to hair follicles via a novel liposome composition.
XX The liposomes are comprised of a lipophilic or lipophobic compound which
XX will selectively target the hair follicle (via the stratum corneum)
XX without damaging or unwanted effects on cells outside the follicle
XX Compound e.g. P-glycoprotein, can be delivered to the hair follicle for
XX treatment of chemotherapy-induced alopecia. Other compounds targeted at
XX hair follicles can include: tyrosinase (or the DNA encoding it -
XX AAO72871), melanin or hair dyes (to restore hair colour or condition);
XX human transforming growth factor-alpha (AAO72873) (for reversal of wavy
XX hair); hair growth stimulants (e.g. cyclosporin A or aromatase) or
XX antisense sequences. The method allows compounds (e.g. macromolecules
XX or polymers), which would not normally reach the hair follicles, to be
XX delivered to these target areas.
XX
XX Sequence 4646 BP; 1371 A; 893 C; 1128 G; 1254 T; 0 other:
XX
XX
XX Query Match 33.1%; Score 1225.2; DB 15; Length 4646;
XX Best Local Similarity 60.9%; Pred. No. 0;
XX Matches 2178; Conservative 0; Mismatches 1266; Indels 132; Gaps 6;
OY 162 caggtgacccctgattatgttggaatgagtggtgtcgtccttgatttggtaacata 221
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 760 caggtatgctcattatttaacagtgaatgtgtgtgtgtgtgtgtgtgtcattca 819
OY 222 gatttccctgtgtgattataactgacgacgacgacgaagagagattggaacagtttt 281
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 820 gtttcatttggctgtgctgtgagacaaatatacaaaatagaagcagtttt 879
OY 282 ttattcagtttggacacgacatcggtgtgttatagctgtgacatggtgaactaa 341
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

```

Dh 880 tcatgtctaataatgacagagagatagctggtttgatgtgcacgattgttgggaacttaa 939  
OY 342 cactgcgatgacagacattgac---aaatcagtgatggtatgagabaagtgtct 398  
Dh 940 caccgacttaacagatgctctcaagaattaaatgaattgtgtgcacaaattggaa 999  
OY 399 gtgtgttcaaaacatgctactcttcttgattggtcgtgcagttgtgttgaaggctg 458  
Dh 1000 gtctcttaagtaataaggaaacttttcaatggtgttataagtaagattacacgtgtgtg 1059  
OY 459 gaactccacctagtagactctacacgctctcctcttaataagttcagcgcgacgatg 518  
Dh 1060 gaagtaacacctgtgattgttggccatcagctcgtctctcttgacgtcagcgtcgtctg 1119  
OY 519 ttctagatggttcattcatttgacgaatgaagaaatgaagttccatltccaaagctgggc 578  
Dh 1120 ggcacaaagtactatcttacttaactgaataaagaactctagcgtatgcaaaagctggagc 1179  
OY 579 tgtggcagaagaagctctgtctacatccgaacgctcattagccttagggccagagaa 638  
Dh 1180 agtagctgaagaggtctctggcagcaattagaactgtagatcttggaggaacaaagaa 1239  
OY 639 agaacttcaaaagttcttctctttaaataacaagaatagcttgttatttcccca 698  
Dh 1240 agaac----- 1244  
OY 699 gtgctactaagtgtgttctgttntttgtaagttatcacagaatcccaagatgcacaa 758  
Dh 1245 -----ttgaaggtacacaaaaatttgaagaagctaa 1278  
OY 759 ggaatttgcataaaaaggactatagcttcaaaagtctctgtgtgtgtacttct 818  
Dh 1279 aagaatttggataaagaagcattacagccaataattcttaagttgcgtcttctgtct 1338  
OY 819 tatgaatggaaactatgactgtcttltgtatggaacctctcctgattcttaatgagaa 878  
Dh 1339 gatcatgtacttatgtctcgtcctctcgtatgggaacctctgtcctcctcaaggaa 1398  
OY 879 acctgtgatatcaatcggagactgtctcgtcgttcttctttagtgaatccatagcagtta 938  
Dh 1399 -----atatctatgtgacaagtaactcactgttacttcttctgtatnaattggggcttt 1452  
OY 939 ttgcatttggagcagcagcctccctcacttggaaaccttcgcaatgcccagggagctt 998  
Dh 1453 tagtgttggacagcagcactctcgaagcattggaagcatttgcacaaagagggagcctta 1512  
OY 999 tcaatatttccaggttatattgataagaacccagtaataagaaacttccacagctgtgata 1058  
Dh 1513 tgaactttcaagataattgataataagccaagtattgacagctattcgaaagatgggca 1572  
OY 1059 taaacctgaatccatagaagaagcagctgtgaaatttaaaaaatgttcttccaattatccatc 1118  
Dh 1573 caaacagataatataatlaagggaaatttggaaattcagaatgttcaacttcaatccacatc 1632  
OY 1119 aagaccatcatcaagattctgaagaaggtcctgaatctcagaattaaagcttggagagacgt 1178  
Dh 1633 tcgaaagaagaattgaagacttgaagggccttgaaacctggaagtgcagagtggaacaggt 1692  
OY 1179 cgccttggtcggtcacaatggtcagtggtggaagagtaacgttagtccagcttctcagaagtt 1238  
Dh 1693 ggccttggttggaacagtggtcgttgggaagagcaacacagtcagctgtatgcagaagct 1752  
OY 1239 atatgtatccgatgtatgcttataatgttggatggaatgaatcagacagagctttaaagt 1298  
Dh 1753 ctatgtaccaccaagagggtatgtgtcatgttltgtagacagatatagaccataatgt 1812  
OY 1299 ggcgattatcgagacacatattggagtggttagtcaagaagcctgttgttgggacac 1358  
Dh 1813 aaggttcttaagggaataatcattgtgtgtgtgagtcaggaacacctgttattgttccacac 1872  
OY 1359 catcagtaacataatcaagatltgacgagatgtatgtactgtatgaagatggagagac 1418  
Dh 1873 gatagcttgaacaaatctgctatgtgcggtgaaaaatgtlcaaccatggttgatggaaagc 1932

OY 1419 agcaagggaagcaaatgctgatgtattatcatatgagtttccataaattacaatc 1478  
Dh 1933 tgtcaagaagccaaatgcttatgacttatactgaacatgtcctcataaatttgacacct 1992  
OY 1479 ggttaagggaaggaagcctcaatgaagtggagggaaggaaggaatgtcagttctg 1538  
Dh 1993 ggttggagagaggggccagttgagttgggcgaagcagaagatctgcacttgcag 2052  
OY 1539 tgccttagtgcgaaccccaagaatctgaattttagatgagcgtctcgccttgatc 1598  
Dh 2053 tgccttgttgcgaaccccaagaatcctcctcgtctgtatggagccagcgtccttggacc 2112  
OY 1599 agaaagcaagtcagctgttcaagctgcacgtggaagagcgagcaaaagttcgaactaacat 1658  
Dh 2113 agaaagcgaaagcagtggttcaagtggtcctgtgataagcgcaaaaagttcgacacacat 2172  
OY 1659 cgttggtagacacagcacttctactatcgaagtgcaagtttgattgttgtgacctaaaga 1718  
Dh 2173 tgtgatagtctatcgttctgtccacagttcgttaatgtctgcgtccatcogctgttgcata 2232  
OY 1719 tggaaatgtgcggaaggaagacatgtctgaactaaatgycacaaacgagttalatba 1778  
Dh 2233 tggagtcattgttggaggaaggaatcatgatagtactcatgaaggaaggaagcattact 2292  
OY 1779 ttcacttgtatgtcatcagagatatataaaagctatgaacagatvggagtaataata 1838  
Dh 2293 caaacttgcacaatgacagcagcaagatgaattgattatgaanaatgacagctgata 2352  
OY 1839 tctta-----ctgaagaagaagacaaactcctccttgcactctgtg 1881  
Dh 2353 atccaaagtgaatattgactgtccttggaaatgtcttcaaatgtatcaagatccagctaat 2412  
OY 1882 aagagcatca-----agtcagacttcatgaaaggtcgtgagga 1919  
Dh 2413 aagaaagaatcatcaactcgttaggagtgctcgtgatcaacagccagacagaagaacttag 2472  
OY 1920 atccacccaatataagaagataagctctccctgaagttctctcattaaaaattttaagtt 1979  
Dh 2473 taccaaagaggtctcgtgataagatactccacatlttcccttltggagagatlatgaagct 2532  
OY 1980 aacaagcctgatatggccttctgttctgtggagcattggtcttcgtctcaatggaac 2039  
Dh 2533 aaatttaactgaatvgccttatttgtgtgtgtgtatttltgtccatataataatggag 2592  
OY 2040 tgttcatcagatlttccatcatccttgcacaaataatlaaacatgtt---ggaatlaa 2096  
Dh 2593 cctgcacacagcatttgcataataatttcaagaatttagggcttttaacaagaattga 2652  
OY 2097 tgaataaacacatataagcagatgacgaagaattatctcatgataatcogtaatttgg 2156  
Dh 2653 tgaatcctgaacaaacagacagaatgaatactgttlttcaatgttcttgccttgg 2712  
OY 2157 tgtatttgccttgtcagttatttcatgaggaattattttagcggacagaggggaat 2216  
Dh 2713 aattatttcttattataatttctcctcagggtttcaaatlttggcaaacgctggagagat 2772  
OY 2217 tttaacgatgaatlaagacacttggccttcaagccatgltatalatcaggaatattgtcgt 2276  
Dh 2773 cctcaacaaagcgtcccgatataatgttttccgacatcagtcagacagagatgtagt 2832  
OY 2277 gtttgaagaaaggaacagcagagaggttgaacacaaatataagcatatgatatagc 2336  
Dh 2833 gtttgaatgcccttaaaacacacatggtgacattgtctacacaggtctgcgaatgatatg 2892  
OY 2337 acaaatcaaggaaggaacaggttccagagattgcttcaacaaacaaatgcaactaacat 2396  
Dh 2893 tcaagttaaaaggtctataagttccagagctgtcgttaattaccagaatataagaaatc 2952  
OY 2397 ggaacttcaagttacattccttataatataatgagtggaagatgacattcctgtatctgag 2456  
Dh 2953 tgggacaggaataatataatcctcatcatalgttggcgaactaacaacatgttactcttagc 3012

QY	2457	tattgtccagagtaattgtccgttgacaggaatgtgttaaacccgagcaatgactgtattgtc	2516
Db	3013	aattgtcccccatcattgtcaaatagcaggaattgtgttaaatgtaaatgtgtctgttgcgaagc	3072
QY	2517	caacaagaataaagcaagaacttaagacatgtcttgaaagaatagcaactgtgaagctttgtgaa	2576
Db	3073	actgaaagaataaagaagaactagaaggtgtcttggaagatcgctactgtgaagaataagaaa	3132
QY	2577	tatacgtactatagtgtcatttaacaaagggaaaaagccttcgagcaaatgtatgtgaagat	2636
Db	3133	cttcgcgaaccggttgttcttctgtgactcgaagagacagaagattgtgaacatagtatgtctcagag	3192
QY	2637	gcttcacacccaacaacaagaataatcccgcgaagaagcacaagatattgtgaagactgtttgtc	2696
Db	3193	tttgcagtgactccatacagaaacctcttttggaagaagacacactttgttaatactttc	3252
QY	2697	attcagccatgtcccttatataatttttcctcatabcagcaaggtttcgtactttgagcctatt	2756
Db	3253	cttcaccccaagcaatgatgtatttttccatgtctgagtttccggttttgagcctactt	3312
QY	2757	aattcaagctgtgacgaatgtgaccccaaggagcagtltcatagttttactgtcaattgtcata	2816
Db	3313	gttgtcacataaactacatgagcttttgaagatgtctcgttagtattttacagctgtgtctt	3372
QY	2817	tggagctatgtgcgatctcggaanaaacgcttcgttttgctcctgtgatatcttaaacgccaatc	2876
Db	3373	tgtgtccacatgtgcgttgagggaagtcagatcatttgccttcgcctatgtccaaagcccaat	3432
QY	2877	gggggtcgtgcacatcgttttccctgttttgaaaaaagaaaccaatatagacagccgcatca	2936
Db	3433	atcagcagccccaactatcatctgtatctgttaaaaaaaccccttgtattgtgcacataagac	3492
QY	2937	agaaagcgaaaagccagacacacatgtgaagggaaattagaagtttcgagaagctcttctt	2996
Db	3493	ggaagcgctaatgtccgacacacatltggaaggaattgcacatttgttgaagtgtatctaa	3552
QY	2997	ctatccatgtgtccgccagatgttttctactctcgtgtgctatctccccaatatttgagcagag	3056
Db	3553	ctatcccaaccgagccgagacatctccactgttctcaaggtactcgaagccgttgaggtgaagaaagg	3612
QY	3057	aaagacagtagcatgtttgttgaggagcgcgtgttggaagaacactctgttccaactctt	3116
Db	3613	ccagagcgtgcctcgtgttggtgcagcaggtgcgttggaagacacaggtgtccagctctt	3672
QY	3117	gcagagaccttatagacccgtgtgcagaagacacatgtcgtttgtgttggtgtgtgacatgcataa	3176
Db	3673	ggagcggttctactgaccccttgcgacaggaaggtgcgtctgtgtgcgaagaataaacgg	3732
QY	3177	attgaatgtatcagtggtgtcctgtccccaatatagacatcgttcttcaagggctgtgtctt	3236
Db	3733	actgaatgttctcagtggtgtccgagacaacacctgtgtcactcgtgtcccgagggcccatctgtt	3792
QY	3237	caactgcagcatgtgtctagagacaatcgactatgtgtgaacacagccgtgtgttcactaga	3296
Db	3793	tgactgcagcatgtgtgtgagacacttgtcctatgtgagacaacagccggtgtgtgttcacagga	3852
QY	3297	tgagatcaaaagagccgcaaatgtcagcaaatatccatcttattatgaaggtctccctga	3356
Db	3853	agagatcgttgaggcgcaagcaaaaggagccaacatacatgtcctcatcaggtcactctctaa	3912
QY	3357	gaaatacaacaacaagttgtgactgaaaggagacacgcttcttgccggtccgaagaaacaag	3416
Db	3913	taaatatagactaaatgaagagacaaaggaactcagctctctctgtgttgccagaaacaag	3972
QY	3417	actagcatgttgcagggctcttccaaaaaacccaatatcttatgttggatgtgaagccac	3476
Db	3973	catgtccataatcgtgcgtctgtgttagacgcgtccatctatttgcgtttgtgaagaagccac	4032
QY	3477	ttcagccctcgtatatatagcgttgagaaggtgtgttcagcagctgccttgataaagccagagac	3536
Db	4033	gtcagctctgtgatatcagaaggtgaaaaagttgttccaaagaagccctgtgcacaagccagga	4092
QY	3537	gggaagagacatgctagtgtgtcaccacaggtcctctgcaatttcaggaacgcagattgtat	3596

[illegible]

RESULT	8
AAZ94738	
ID	AAZ94738 standard; cDNA; 4646 BP

DT 01-AUG-2000 (first entry)

DE Human ATP binding cassette ABCB1 (MDR1) cDNA

KW ABCB1, ATP binding cassette, human; cholesterol; lipid disorder;  
 KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;  
 KW lupus erythematosus; diagnosis; gene therapy; MDR1;  
 KW multidrug resistance; chromosome 7q21; ss.

OS Homo sapiens

PN W0200018912-A2

PD 06-APR-2000.

PF 21-SEP-1999; 99WO-EP06991.

PR 25-SEP-1998; 98US-0101706.

PA (FARB ) BAYER AG.

PI Schmitz G, Klucken J;

DR WPI: 2000-293151/25.

**PT Adenosine triphosphate binding proteins useful for identifying agents**

PS Claim 9; Page 110-112; 154pp; English.

CC The present sequence is that of human ATP binding cassette  
CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a  
CC differential display method in which monocytes from peripheral  
CC blood were subjected to macrophage differentiation and cholesterol  
CC loading with acetylated low density lipoproteins and subsequent  
CC cleaving with high density lipoprotein (HDL3) to identify  
CC cholesterol sensitive genes. The gene maps to chromosome 7q21  
CC and is also termed MRL (multidrug resistance). The invention  
CC provides cholesterol-sensitive ABC genes (see AA924374-63). These  
CC genes, and polypeptides encoded by them, can be used for diagnostic  
CC and therapeutic applications, and for biochemical or cell-based  
CC assays to screen for pharmacologically active modulator compounds  
CC useful for the treatment of lipid disorders, atherosclerosis or  
CC other inflammatory diseases such as psoriasis and lupus  
CC erythematosus.

Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Query Match	33.18;	Score 1225.2;	DB 21;	Length 4646;
Best Local Similarity	60.98;	Pred. No. 0;		
Matches 2178; Conservative	0;	Mismatches 1268;	Indels 132;	Gaps 6

OY 162 cagattgacctgtatcatcgttgaataaggctgcgccttgattttggttacataca 22H  
||||| || ||||| ||||| ||||| || ||||| ||

Dh 760 caggtatgacctatattlaacaagtgaattggtgctgggtgctggttgcgttaccatcca 819  
Qy 222 gatttccttggtgattataactgcagcgacgacgacgaaggagatccgaataattttt 281  
Dh 820 ggttcattcttggtgctccgagcgtggaagacaatatcacaaattgaataaacattttt 879  
Qy 282 tcatcaatttttggcacagagacatcgtgctgtttgatagcgtgtgacacggtgaaatca 341  
Dh 880 tcatgctataatgacagcgagagatagctggtttgattgtgcagcagtgttgggagattca 939  
Qy 342 cactgcgacgacagacattgc--aaatacagtatggtattggagataaagattgctct 398  
Dh 940 caccgacttaacagatgltcgtcctaagaattaatgtaattgttggcaaaattggat 999  
Qy 399 gttgtttcaaaaactgttacttttttcgattggcctgagctgagttggttgggaaaggctg 458  
Dh 1000 gttcctttagtcaatggaacaatttttcaactggtgttaataagattacacgtgtgtg 1059  
Qy 459 gaaactacacctagtgactatccacgtccctccttaataatggttcagcgacagcagtg 518  
Dh 1060 gaagctaacaccttgttattgttggccatcagtcctgtcttggactgtcagctgctgtctg 1119  
Qy 519 ttctagaatgtgtcattcatttgcacgaataaggaattaaagtgcctattccaaagctgggc 578  
Dh 1120 ggcacaagtactatccttacttacttgcataaagaactccttagctatgcaaaagctggagc 1179  
Qy 579 tgttgcaagaagaagtctctgttcaatccgacgaagtcacgtccttaaggcccgagagaa 638  
Dh 1180 agtgcgtgaagaaggtctctgtagcagcatatgacactgtgattgcatttggagacaaagaa 1239  
Qy 639 agaacttcaaaagctcttccctttaataatacaagatatgctgtgtttatttcccca 698  
Dh 1240 agaac----- 1244  
Qy 699 gtggtactaaagtgtgttctgttnttgtlaagytatacacagaatcacaagatgycaaa 758  
Dh 1245 -----ttgaagytacaacaanaatttagaagaagctcaa 1278  
Qy 759 ggattttggcctaaagaagactataagctcacaagaagtcctctgtgtgctgtgactctct 818  
Dh 1279 aagaatttggataaagaagactatacagccaatttctctaaggtgtgctgtcttccgt 1338  
Qy 819 tatgaatggaacactatgactgtccttggtagaacccttgattccttaatcgggaga 878  
Dh 1339 gatctatgacacttatgtctcgtgctctgtgtatgagacacctgtgctccctcaggagga 1398  
Qy 879 aacttgatataccatcggagactgttctbgtcgttttctttagtgaatccataagcagta 938  
Dh 1399 -----atatctatltggaacaagtaactccttcttctgtatatttggtgcttt 1452  
Qy 939 ttgcatgtggagcagcagtcctcctcacttgaaccttcgcaatagccgagagctgctct 998  
Dh 1453 tagtgttggacaagcaltccaaagcatltagaagcatttgcgaatgcaagagagcagctta 1512  
Qy 999 tcatattttcagagttatagtaagaacccaagratatataacttttccacagcgtgata 1058  
Dh 1513 tgaactctccagaataatctgaataaagccaagatltgacgctatttgcgaagagtg99ga 1572  
Qy 1059 taacactgaatccatagagaagactgtgaaatttaaaaatttttcttcaattacatc 1118  
Dh 1573 caaacccagataatattgaaggaaatttggaaatccagaatttccactcagttaccctc 1632  
Qy 1119 aagacactatcaagatctgaaaggtctgaatctcgaattaaagltctgtagagacagct 1178  
Dh 1633 tcgaanaaagaattgaagcttgaagggccttgaacctgaaggtgcagagtgtagcagagcgt 1692  
Qy 1179 cgccttggtggtcctcaatgcagtgtaggaagagtagcgttagtccagcgtctgcgaaggtt 1238  
Dh 1693 ggccttggttggaaacagtggtctgtgggaagagacaaacgtcccgatgtagcgaaggtc 1752  
Qy 1239 atatgataccgagatgagcttcatatgltgtatgtagaatacatcagagctttaaagt 1298  
Dh 1753 ctatgaccccaagagggagtgatggtcagtgtgtagacagagatatlaagaccataaagt 1812

Qy 1299 gcggtattatcgagacatatttggagttggttagtcaagaagcctgtttgttcgggacac 1358  
Dh 1813 aaggttctcagggaaatcatctgtgtgtgtagtcaagaaacctgtattgttgcacac 1872  
Qy 1359 catcagtaacaaatlaacaaatgatgagcagatgatgtgactgataagagatgtagagagc 1418  
Dh 1873 gatagctgaanaacatctgctatgctgcgcgtaaaatgtccacatgagtgtagagaaagc 1932  
Qy 1419 agcaagggaagcaaatgctglatatlttaatactgaggttccctaataatlaatacatt 1478  
Dh 1933 tgcagaaggaagccaattgctatgacttatactgaactgctcctcataaatttgaaccc 1992  
Qy 1479 ggtlaagggaanaagagctcaaatatgagtggagggaggaagaaacgaagagatgcagttgcg 1538  
Dh 1993 ggttggagagagaaggccagcttgagtgtgtgtgtagaagacgagagatccgcatctgcag 2052  
Qy 1539 tgccttagtgcgaacccccaagattctgatttlaagatagcgtacgtcctgcctgattc 1598  
Dh 2053 tgccttggttcgcaacccccaagatctcctcgtctgtgataagcgcaagctcagccttgagac 2112  
Qy 1599 agaaagcaagtcagctgttcaagctgcactgtagaaggcgagcaaaaggtcggactacaat 1658  
Dh 2113 agaaagcgagcagtggttcaagtgctgtagatgaagcgcaagaaaggtcggacacat 2172  
Qy 1659 cgtgttagcacaacgacttcttactatctcgaagtgcagatttgaattgtgacccctaagga 1718  
Dh 2173 tgtgaatgctacgtcttgtctacgtctgtaattgtatgtacgtcactgctgtgttctgaaga 2232  
Qy 1719 tggaaatgctggcggaaggaagcgacatgcgtgaactaattgcaaaaagagcttatla 1778  
Dh 2233 tggagctaatgtggaggaaggaataca tgaatgaatcattgaaggaaggaagcatttact 2292  
Qy 1779 ttcaactgtgatgtcacaggaatataaagctgtagaagatgtagatcacatga 1838  
Dh 2293 caaactgttcaaatgacagacgaggaatgaattgaattagaataatgcagctgtaga 2352  
Qy 1839 ttctta-----ctgaaggaagaccactcactccctgcagctgtg 1881  
Dh 2353 atccaaaagtgaatttgaatgcttggaaatgtcttcaaatgtattaaagatcagaactat 2412  
Qy 1882 aagagcatca-----agtcagactcatcttgaacagcgtgagga 1919  
Dh 2413 aagaaagaatgaatcgttagaggtgtccgtgtgatacacaagcccaagacagaagaacttag 2472  
Qy 1920 atccaccaatctaaagataagtagtcttcccttgaaggtctctctatataaaattttaaagt 1979  
Dh 2473 taccaaagaggtcgtgtgataagataacccctcaggttcccttggagagattagaagct 2532  
Qy 1980 aacaagcctgaatggtccttgtgtgtctcgtgggaacattggtcctctgttctaattggaac 2039  
Dh 2533 aaatttaactgaatgtccttatttgtgtgtgtgtgtgttatttltgtccattataaagagga 2592  
Qy 2040 tgttcatcaggtatttcatcatcttgcgaanaattataaacacgttt---ggaataaa 2096  
Dh 2593 cctgcaaccagcatttgcataataatlttcaaaagttabaagggttlttacaagaagtga 2652  
Qy 2097 tgaataaacacatlaaagcattgacgagaaatttatttcatgataatcgttcaatttgg 2156  
Dh 2653 tgaatcctgaacaaacgaacgaagaatagtaactgttccactatgttcttccgcttgg 2712  
Qy 2157 tgtatttgccttgtcagttatlttcaatgcaggtatttattgaagcagagcagggagaaat 2216  
Dh 2713 aatttcttcttattatatttlttccctcaggggttccatttgcatttggcaagcgtgagagat 2772  
Qy 2217 tttaacgattgaatlaagacacttgccttcaagcactgataatatacaggaatgacctg 2276  
Dh 2773 cctcaacagcggctccgatataggttltccgataccatgctcagaacgagatgtgagtgt 2832  
Qy 2277 gtttgaatgaaggaacacacagcagagaggttgcgaacaaatatatgacatagatatagc 2336  
Dh 2833 gtttgaatgaccttaaaaacacacatcgagagcattgtgactaccagcgtcgcgaatgagctgc 2892

QY	2337	acacattcaaggagcagaaggttcccaaggttggcgcttcctaacacaaaatgcaatacat	2396
Db	2893	tcaagttcaaaaggggccatagttcaggttcgcgtctgtaattccccaagatataagaatact	2952
QY	2397	ggagcttcagttatcatcttccttataatgatgatggagatgacatccgatctccag	2456
Db	2953	tgggacagagatattatatacttccttcatactgatgttggcaactaaacctgttaccttagc	3012
QY	2457	latgtctccagtaactctgcgttgacaggaatgatgtgaaaacccgacgaatgacgtattgc	2516
Db	3013	aattgtaccatcatctgcaataagcaggaattgtgaaatgaaaatggtctgctgacaagc	3072
QY	2517	caacaaagataaagcaagaacttaagacatgctgtgaaaagatagcaactgtgaagcttgggaa	2576
Db	3073	actgtaaaagataaagaagactagaagtgctgtgggaagatcgtctactgtgaagcaatagaanaa	3132
QY	2577	latacgtacataatgtctcatatlaacaaggaaaagcccttcagcacaatgatagtgaagat	2636
Db	3133	cttcgcgaaccggttgttcttcttgacttcagagacagaagattgaaacatatgtatgctcagag	3192
QY	2637	gcttcagacacccaacacagaataatccctcgcgaagaaaagcacagataattatgaaagctgtatgc	2696
Db	3193	tttgcagtagtaccatacagaanaactctttggggaagcacacatcctttgaaattacatttc	3252
QY	2697	atccaagcattgcttatataattttgctcctatgcagcaggggttcgatttgaacctattc	2756
Db	3253	cttccacaaaggacaatgatgtatatttttccctatgctgtgatgttccgggttttggagcttactt	3312
QY	2757	aattcaagcctggagcagatgacccccaagaggcagtgctcatagatttttaactgaattgacata	2816
Db	3313	ggtgtgcacataaactcgaatgcagcttttgaggaaigtctcgttagattttcagctgtgttcctt	3372
QY	2817	tggagctatgcccacatcggaanaaacgctcgttcttgctccctgaataattccaaagccaaatc	2876
Db	3373	tgtgtgcattgcccgttgggggaagtcagattcatcttgctcgtgactgacaaagccaaat	3432
QY	2877	gggggctgcgcgcatctgttttgccttgcttgtaaaaaagaaccaaatataagcacgcgagta	2936
Db	3433	atccgcagccccaacatcatcatgatactgataaaaaaccccttgatgtacagctacagac	3492
QY	2937	agaaagggaanaaagccagacacacatgctgtaaagggaattagaagtttcgagagctcttcctt	2996
Db	3493	ggaaagccctaatgccgaacacacattggaagggaatgatgcaatttggtaagttgattccaa	3552
QY	2997	ctatccatgctccacagatgttcttcatctccgttggtcctatccctcagtatgtagcgagag	3056
Db	3553	ctatcccaaccgacccgacatccacagctgctcttcagggacgtcagccctcgagggtgaaagagg	3612
QY	3057	aaagaaacagtatgcatcttgctgggagacagcgctgtgggaaagacatctctgtcaactctc	3116
Db	3613	ccgaagcctgctcgtctgtgtggcgagcagctgtgcgttgaggaaagacaagtggtccagctctc	3672
QY	3117	gcagagagccttatgaccccgctgcaaggaacaaagtgcgttttagtggtgtgggtgcaaaaga	3176
Db	3673	ggagccggttctcaacacccctctggcagggaaagtgctgcgtctgtatggcgaagaataaagcg	3732
QY	3177	attgaatgatacagtggtgcctccgttccccaataagcaatcgtcttcacagaagcctgtgtcctt	3236
Db	3733	actgatagtctcagtgctccgagcacacctgycgcatcgtgtccagagggcccatctcgtt	3792
QY	3237	caactgcagcatctgtctgagaacatcgctctatggtgacaacagcgtgtgtgtgcattaga	3296
Db	3793	tgactgcagcatctgtctgagaacattgctctatgtagacacaacagccgggtgtgtgcacaaga	3852
QY	3297	tgagatcaagaagacgcgcaaatgcaagcaatccatcttttatgttgaagttccctcgga	3356
Db	3853	agagatctgtgagggccagcaaaaggccacacatacattgcttcatcgcagtcacgtccaa	3912
QY	3357	gaataacaacaacaatgttgactgtgaaagagacagcgttcttggcggccggaanaaag	3416
Db	3913	taaatatagccactaaagttagagacaagaaggaactcagctcctctgtgtggccggaanaaagc	3972
QY	3417	actagctatttgcaaggcctcttctccaaaaaaccaaatatttatgttggatgagccac	3476

Db	3973	catgtccatagctcgtgcctctgttagacagccatatttgccttggatgaagccac	40332
QY	3477	tttagcctctgataaagaaagtgtagaaggtgtttagacgtcccttgaagcagagac	3536
Db	4033	gtccagccctgtgatacagaagaagtgaaggttttccaagaagccctggacaagccagaga	4052
QY	3537	gggaagaacatgcctagtgtgtccatccacagctctctgcattctagaagcagatttgat	3556
Db	4093	agagccgacactgcatltgtattgtcttcaacgcctgtccacatccagatgtgaactaat	4152
QY	3597	agttgtcttcgacaatgtgaagaagaataaaggacaagaactcatcaagactcctctggaaa	3656
Db	4153	agtggtgttccagaatggtgcagagtcgaagatgtgcacgtcatcagcagctgtgtgcaca	4212
QY	3657	tcgagacatatatttaagttaatgtatcacaagtcag	3694
Db	4213	gaaagcactcatatttccaatgtgcaggtgtccaggtg	4250

XX	RESULT	9	
XX	ID	AA113394	
XX		AA113394 standard; DNA: 6505 BP.	
XX	AC	AA113394;	
XX	DT	24-JUN-1996 (first entry)	
XX	DE	Hybrid vector pSF-MDR.	
XX	XX		
XX	XX	Hybrid; vector; gene transfer; gene therapy; haematopoietic stem	
XX	XX	retroviral; murine embryonic stem cell virus; MESV;	
XX	XX	MoJoney murine sarcoma virus; (MOMUSV);	
XX	XX	Friend murine leukemia virus; F-MuLV; ds.	
XX	XX	Synthetic.	
XX	XX	DE19503952-A1.	
XX	XX	14-MAR-1996.	
XX	XX		
XX	XX	07-FEB-1995; 95DE-1003952.	
XX	XX	08-SEP-1994; 94DE-1413973.	
XX	XX	(BOEF ) BOEHRINGER MANNHEIM GMBH.	
XX	XX		
XX	XX	Baum C, Ostertag W, Stocking-hardbers C, Stockinghardbers C;	
XX	XX	WPI; 1996-152306/16.	
XX	XX		
XX	XX	Hybrid retroviral vectors - for gene transfer into haematopoietic	
XX	XX	stem cells	
XX	XX		
XX	XX	Disclosure; Page 25-29; 42pp; German.	
XX	XX		
XX	XX	New hybrid vectors comprise (1) a leader region including the U5	
XX	XX	region and tRNA primer binding site of murine embryonic stem cell	
XX	XX	virus (MESV) or MoJoney murine sarcoma virus (MOMUSV), and (2) a	
XX	XX	LTR including the U3 and R regions of a Friend murine leukaemia	
XX	XX	virus (F-MuLV). The vectors are useful for ex-vivo or in-vivo g	
XX	XX	therapy. High levels of gene transfer can be achieved in	
XX	XX	haematopoietic stem cells and their myeloid (non-lymphatic) prog	
XX	XX	PSF1, PSF2, PSF3 and PMM1 (sequences given in AA113390-T13393) a	
XX	XX	examples of such vectors.	
XX	XX	Vector pSF-MDR (sequence given in AA113394) is based on the	
XX	XX	MESV vector R224.	
XX	XX		
XX	XX	Sequence 6505 BP; 1719 A; 1506 C; 1606 G; 1674 T; 0 other;	
XX	XX		
XX	XX	Query Match 33.18; Score 1225.2; DB 17; Length 6505;	
XX	XX	Best Local Similarity 60.98; Ptd. No. 0;	

Matches 2178; Conservative 0; Mismatches 1268; Indels 132; Gaps 6;

```
QY 162 caggttgacctgtattatgttgaaatggtgtgtccttgatttttgttatcata 221
    ||||| || ||||| ||||| ||||| || || || ||||| ||
Db 2152 caggtatgccttatcttacaagtgaaattgtgtggtgtcgtgtctatcata 2211
QY 222 gatttcctgttgatataactgcagcgacgacgaacgaagagatctgaaacatttt 281
    ||||| ||||| || || || ||||| || || || ||||| |||||
Db 2212 gtttcaattctgtgtccgtgcagctggaagacaatacaaaattagaacatttct 2271
QY 282 tcatcagttttgacacagacatcgctgtgttatagctgtgacatcggtgaacttaa 341
    ||||| ||||| ||||| ||||| ||||| || || || ||||| ||
Db 2272 tcatgtctataatgcgcacgagagatgctgtgtgttgatgtgcacgagtgttggaagacttaa 2331
QY 342 cactgcgcatgacacatgac---aaaatcagtgatgtatgtgagataaagattgtct 398
    ||||| ||||| || || || ||||| ||||| ||||| ||||| ||
Db 2332 caccgaccttacaagtgtctctcaagaattaatgtatgtgtgacaaatttgaat 2391
QY 399 gtgtgttcaaaaacatgtctactttctgattgtgcctgagctgtgtgtgtaaggctg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2392 gtcttcaagtcaatggaacaatttctcaatggtgttatagtagaattacaagctgtgtg 2451
QY 459 gaacctcaccttagtgactctataccacgtctctctataatgagcttcacgcgcagcatg 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2452 gaagctaacacctgtgtattgttgccatcagctctgttcttgaacgtgcgcgtgtctg 2511
QY 519 ttctagatgtgtcatctcaatttgaacagtaaggaattagtcctattccaaagctgggc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2512 ggcagaagctactatctcaatttaccgtataaagaactcttagcgtatgcataaagctgagc 2571
QY 579 ttgtgcagaagaagctctgtcatatccgaatccagatcagctcttagagccacgaagaa 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2572 agttagctgaagaggtcctgtgcgcagcatatgaacgtgtgattcttgcggagcaaaagaa 2631
QY 639 agaacttcaaaagcttcttccctttaataataacaagatatgctgtttatttcccca 698
    |||||
Db 2632 agaac----- 2636
QY 699 gtgtgctaagtgtgtctgttntttgttaaggtatatacagaactctcaaaagtgcata 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2637 -----ttgaaggtacaacaaaatttgaagaagctaa 2670
QY 759 ggaatttggcataaaagagactatagcttcaaaagtgtctctgtgtgtacttct 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2671 aagaatttgggataaaagaagcatcttacaagccaatttctataggtgtgcgttccctgt 2730
QY 819 tatgatatgaacctatggaactgtcttattgtgtatgaaacctctgattcttaatggaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2731 gatctatgacatttctgtcctgtgccttctgtatgtggacacctgtgtcctctcaaggga 2790
QY 879 acctggaatacacatcggaactgtctctgtgttcttctttagtgttaatccatagcaatga 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2791 -----atacttatgtgacaagctacacgttattcttcttgaattaatctggggcttt 2844
QY 939 ttgcatgtgagagagagctccctcactttgaaaccttgcgaatagcccgagggctgcctt 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2845 tagtgttggacaggaactctcaagcatctgaagcatcttgaatctgaagagagcgactta 2904
QY 999 tcaattttccaggttatattgataagaacccagtatagtaacttttccacagctgtgata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2905 tgaatcttcaagaataattgtataataagccaagtattgcacgattctgaaagctgggca 2964
QY 1059 taacctctgaatcatalagaagaactgttgaaatttaaaaaattgttcttcaattatccalc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2965 caaacacagataaattaaaggaatttgaattcagaatgtctcacttcagttaaccac 3024
QY 1119 aagacacatcatcagaatcttgaagaagtctgaaatctcgaatgaatgaatcgtggagagacat 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3025 tcgaagaagaagcttaagaatcttgaagggttgaaacctggaaggtgcgcagagctgt 3084
QY 1179 cgcctgtgtcgtgtcctaattgcagctgtggaagagtaacgtgtgtcgcagctctcaagaggt 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3085 ggcctgtgttggaacaggtgtgtgtggaagagcaacaagctccagctgtatgtcagaaggt 3144
```

```
QY 1239 atatgcgagatgtagtcttatcatgttgatgtagatgacatgcacagacttaaatgt 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3145 ctatgaccaccaagagggatgtgtcagtgcttgaatggaacgatacttgaagacataaatgt 3204
QY 1299 ggcgcattatcgaaccatattgcagtggtgtgtagtaagaagctgtgttctgcggaccac 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3205 aaggttctcaggggaataatcatctgtgtgtgtgtagtaagaacactgtatgtttggccaac 3264
QY 1359 catcgaacaatatacaagtatgacgagatgaatgtatgacatgagatgagagagc 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3265 gatagctgaanaacatttgcctatgtgcgcgtgaanaattgtcacacatgagtgatggaagac 3324
QY 1419 agcaagggaagcgaatgtgtatgtatttatactatgagtttccataatattcaatcat 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3325 tgtcaagaagaagccaattgctatgacttatactgaacatgctcatataattgaacacct 3384
QY 1479 ggttaggggaagaagagagctcaaatgagtgagggaggaacagagatcgcattgtctg 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3385 ggttgagagagaaggggcccagttgagtggtgtgcgaagcgaagatcgcattgtcacag 3444
QY 1539 tgccttagttcgaaacccccaagattctgattttagatgagctacgtctgccttgatc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3445 tgcctgtgttcgcaaccccaagatcctcctgtgtgattgagggccaagcttgagacac 3504
QY 1599 aagaagcaagtcaagctgttcaagctgcacgtgagagaagggcgaagaagtcggaactcaat 1658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3505 aagaagcgagaagcagtggttcaagtgctctgtatagaagccagaanaagctgcgaacact 3564
QY 1659 cgtgttagcacaccgacttctcactatctgaagtgcagaattgtgtgtacccataagaa 1718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3565 tgtgtagctctcgttctgtctcaacgttgttaattgtctgcgcgtcaltcgcgtgttcgata 3624
QY 1719 tggaaatgtcgcggagaaagagacaaatgctgaaactaatgagcaaaacgagctatatta 1778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3625 tggagtcatgttggaagaaagaaatcatgtatgactcatgaaagaaagaaatcttact 3684
QY 1779 ttcaattgtgatgtcacagagatactaaanaagctgatgacagatgggttcaatgaacata 1838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3685 caaacttgcacaatgacagacagcaggaatgaagtgtgaattagaanaatgcagctgtatga 3744
QY 1839 ttcta-----ctgaagaagaagacaactcactcctctgcactctgtg 1881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3745 atccaaagtgaattgtatgtccttggaaatgtcttcaaatgtttcaagtccaagcttaac 3804
QY 1882 aagagcatca-----agtcaagcttcatgaaacaggtctgagga 1919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3805 aagaaanaagatacaactcgtagagagtgctcgtgatacaacagccaaagacagaagacttag 3864
QY 1920 atccacccaacttaagaagataagctctccgtgaagctctctcaatlaanaaattttaagtt 1979
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3865 taccaaagaggtctcgatgaaagatacctccagtttcttcttggagagatattagaagct 3924
QY 1980 aacaacagctgagatgaccttgggtgtcttgagggaacattgctgtgtcttaatggaac 2039
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3925 aaatttaactgaaatggccttatttgtgtgtgtgtatattgtgcacatataaavggag 3984
QY 2040 tgtcatcagatatttccatcatcctcttgcaaaaataataacaatgttt---ggaataaa 2096
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3985 cctgtgaacagagatttgcataataatttccaagaattatagggtttttaaagaagtga 4044
QY 2097 tgaataaacacataaagaatgacgtgacgaatattatctcatgataatctgcgaatttggg 2156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 4045 tgaatcgtgaanaaagcagaatgtaactgttttcaactatgtttctagcccttgg 4104
QY 2157 tgtatttctgttgcaggtatttcaatgcaaggaattatttcaagcagagcgagggaat 2216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 4105 aattattcttataatcaatttctcctcaaggtttcaaatgttgcgaagctgcggagagat 4164
QY 2217 tttaacgatgagatlaagaacacttggccttcaagccaatgtatatacagaatgtcgtg 2276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 4165 cctcaaccaagcggtcccgataatgttttccgatccatgcacgaagcagatgtgagtgtg 4224
```

```

QY 2277 gtttgatgaaagaaacacgacaggaagcttgacacaatattagccatagatagc 2336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4225 gtttgatgacccataaaacacacgtgagcatgactaccagagctcgcaatgctgc 4284
QY 2337 acaaatcaaggagcaacaggtccagattgagcttcaacaanaatgcaatacat 2396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4205 tcaagttaaagggtctaaagttccaggtctgctgtaattacccaagatataagaactc 4344
QY 2397 gggaacttcgtatcatcttcccttataatataatggaatggagaatgacatccgtatccgag 2456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4345 tgggacaggaataattatlatccctcatctatgcttgagcaactaactgtaactcttagc 4404
QY 2457 tatgtcccaagtacttgcgttgaaaggaatgataatgaaacgcgacgaatgactgtaattgc 2516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4405 aattgtcccatcatcttgcaatagcagaggtgtgtgaaatgtaaaatgtgtgtggaacagc 4464
QY 2517 caacaagaataagcaagaacttaagcagctgcggaagaatgacatgcaactgtaagcttgagaa 2576
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4465 actgaaagataagaaagaacagaaggtgcggaagaatgcctactgaaagcaatagaaaa 4524
QY 2577 tatacgtactatagtgcatataaacaagaagacccctcgagcaaatgtataagaagat 2636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4525 ctccggaacccgtgttctcttgaactcagagcagagaatttgaaacatatgtatgtcagag 4584
QY 2637 gcttcagactcaacaagaataatcctcgaaagaagacagatatgtgaagctgtatgc 2696
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4585 ttgcaagtaaccatacagaactcttggaaagcacacactcttggaaatlatatcttcc 4644
QY 2697 attcagcaagccttataatttgctatgacaggggttctgattgaaagctattc 2756
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4645 cttaaccaaggaatgatatcttccctaaagctggaatgttccggttctggaagcctactc 4704
QY 2757 aattcaagctggaatgaccccaagagcagatgtcatcagtttctactgaaatgata 2816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4705 ggtgacacataaactatgagcttggagatgtctgttgatttccagctgtgtcct 4764
QY 2817 tgaagctatgacatcggaanaacgtcgttctgctcctgaatattccaaagcaaatc 2876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4765 tgtgtcatgtgcgttggggcagtcagtcacttctccttactactccaaagccaaat 4824
QY 2877 gggggctgagcactgttctgctgttgaaagaacccaatataagacgcagtcac 2936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4825 atcagacagccacatcatcatgataatgaaaccccttcttctgacacatacagac 4884
QY 2937 agaaaggaaanaagcagacacatgtgaagggaattgaagtctcgaagaagctcttct 2996
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4885 ggaagcctaattgcgaacacatgtgaagggaatgtcacatttgttgaagtgtatcaa 4944
QY 2997 ctatccatgtccccaatgtttcatcctcgtgtgcttaccctcaagtattgacgaag 3056
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4945 ctatccaccccgacgcgacatcccaagtgctcaggtgacgtgaggtgaagaagg 5004
QY 3057 aaagacagtaagcattgttgaggagcagcgtgtggaagaagcctctgttcaacttct 3116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5005 ccgaagcctggtctgtgtggcagcagtgctgtgggaagacacagtgtgccagcttct 5064
QY 3117 gcagagacttatgaccccggtgcaagacaaatgtctgttgaatgtgtgaaaga 3176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5065 ggaagcgttctcagcaccctgtgcaaggaaatgtctgttgaatgcaagaataaagcg 5124
QY 3177 attgaatgataagctgtcgtcgttcccaaatagcaatcgttcccaagaagcctgtgctct 3236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5125 actgaatgttcaagtggtcccgagacacactgtgcatctgtgccagagcccatctgtt 5184
QY 3237 caactgacagcttctgagaacatcgcctatggtgacaagacgcgtgtgggcatattaga 3296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5185 tgaatgagcagctgtcgtgagaacatctgctatggaacaacagcgggtgtgtacagga 5244
QY 3297 tgaatcaanaagaacgcgaatgcaagaatcatcatcttatttgaagtctcccca 3356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5245 agaagatcgttggggcagcaagaaggccaacatacatgtccttcatgtaactgctcaa 5304
QY 3357 gaataacaacacaaagttgactgaaaggacacagcttctgtcggcgacgaatacaag 3416

```

```

Db 5305 taaatagacactaaagtagagacaagaagaaactcgtctgtgtggccaagaacaag 5364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3417 actagcttgcgaaggctctcttccaaaacccaatatttattgttgatgagccac 3476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5365 catgtccatagctcgtcctctgttgaacagcccatatttcttcttgatgaagccac 5424
QY 3477 ttcagccctcgataatgacagtgaagggtgtgtcagcatgcccctgataaagccaagac 3536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5425 gtcagccttgatcacagaagtgaaagggtgttcacagaagccctggaacaagccaagga 5484
QY 3537 gggaagagacatgcttagtgctacacagcctctcgtcgaatccagaagcagatttgat 3596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5485 aggcgcacccgtcgtgtgactgtctcacgcctgtccacatccagaatgcagaactaat 5544
QY 3597 agtgtctgcacatggaagaataaagaacaagaagaaactcaagcgtcctgagaaa 3656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5545 agtgtgttccagatgacgagtcagagcaaggcagcagcatcgaatcgaagctgtcgaca 5604
QY 3657 tcgagacatatatttaagttagtgaatgcacagtcag 3694
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5605 gaaagcactatatttccaatgtcaggtgcaggtcg 5642

RESULT 10
AA052726 standard; cDNA; 4669 BP.
ID AA052726
XX
AC AA052726;
XX
DF 24-JUN-1994 (first entry)
XX
DE Sequence of human multi-drug resistant gene mdrl.
XX
KM P-glycoprotein; multi-drug resistant gene; mdrl; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 424..4267
FT FT /*tag= a
XX
PN W09324613-A.
XX
PD 09-DEC-1993.
XX
PF 14-MAY-1993; 93W0-US04707.
XX
PR 22-MAY-1992; 92US-0887712.
XX
PA (GENE-) GENETIC THERAPY INC.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Mcdonagh KT, Nienhuis A, Tolstoshev P;
DR WPI; 1993-405805/50.
DR P-PSDB; AAR44297.
XX
PT DNA or RNA sequence for human multi-drug resistant gene MDRL -
PT encodes P-glycoprotein with altered splice site, used in cancer
PT therapy
XX
PS Example; Fig 4; 64bp; English.
XX
CC pMDR2000 contains an mdrl cDNA sequence (AA052726) described in PCT
CC application no. W087/0943m wherein the first 282 bp of the 5' UTR
CC and the last 23 bp of the 3' UTR of the cDNA sequence have been
CC removed. The codon YTN is unclear in the spec., and encodes the AA
CC Leu.
SO Sequence 4669 BP; 1394 A; 893 C; 1126 G; 1254 T; 2 other;

```







Dh	2773	ccatcaacagcgcgctccgataacgtgtttccgatcccatgctccagacagagatgtgagttg	2832
Qy	2277	gtttcgtatgaataaagaaacagacagagaggtcttgacaacaatbatagcccatagatataagc	2336
Dh	2833	gtttcgtatgaataaataaacaacacactgagcatcttgctctaccagctgcgcatagtatgtcgc	2892
Qy	2337	acaaattcaaggagcaacaggttccagaggttggcgtcttaacacaataatgcaatacat	2396
Dh	2893	tcaagatataaaggggctataagttccacaggtctgcgttaataccagcaatatagacaatact	2952
Qy	2397	ggagcttcagttatccattccctcttatatgatgatggagagatagacatcccgattctccag	2456
Dh	2953	tgggacaagcaataataataatccctccatctatggttggcaactaaacactgtaactcttagc	3012
Qy	2457	tattgctccagtaactgycgcgtgacaagaaatgaatgaacaacgagcaatgactgtattgac	2516
Dh	3013	aattgtaccccatctatgcaatagcagaggttgttgtaaatgaataatgtgtctgcggacaagc	3072
Qy	2517	caacaaagataagcaagaacttaagaacatgctcggaaagatagacaatgaaactttggagaa	2576
Dh	3073	actgaagaataaagaaagaaactaagaaggtcgtggaaagatcgtactcgtgaagcaatagaaa	3132
Qy	2577	tatagctactatgaatgycataactaaacagggaaaagccttcggacaatgcttatgaaagat	2636
Dh	3133	cttcggaaacggtgttcttcttgactcgaagagcagaagatttgacaataatgatatgctcaga	3192
Qy	2637	gcttcaagactaaacaacagaaataactcctcgaagaagaacagacagatatttgaaagctgtatgc	2696
Dh	3193	tttcgagtaacatatacagaagaactctttggaaagacacatctttgaaattacatttcc	3252
Qy	2697	attcagacatgaccttatatatatttgcctatgacagaggttctgattttgaagcctattc	2756
Dh	3253	cttcaacccaggaactatgataatttcttccatgctcgtgatagttcccggtcttggagcctaact	3312
Qy	2757	aattcaagcttgagaaatggccccagagagcagatgttcatagtttttacttgcgaattgacata	2816
Dh	3313	ggttgacacataaaytmatggtcctttgaagatgtctgttaagatatttcaagctgtgtgctct	3372
Qy	2817	tgagagctatggcccatctcgaaaanaacgctcgtcttggctctccctgataataatccaaagcccaatc	2876
Dh	3373	tggtgcacatggccgttgggggaagtgcatgcatcttgctctgactatgccaagccaatat	3432
Qy	2877	gggggctgagcatctgttttgccttgttggaaaaagaaaccaatatagacagccgagtca	2936
Dh	3433	atcgcagccccaacatctcatatgcatatgaaaanaaccccttgatctgacgtactacaagcac	3492
Qy	2937	agaagagaaaanaagccagacacatgtgaagggaattagagttcttgagaagatccttctct	2996
Dh	3493	ggaaagccttaatgctcgacaacatctggaaggaataatgccaatttggatgaaatgttatcaa	3552
Qy	2997	ctatccatctgccccaaatgatcttcatccctcccgctgagctatccctccagatattgagcgag	3056
Dh	3553	ctatcccaacccgacccgacatcccatgctctcaaggagacatgagccgtggaggtgaagaaggt	3612
Qy	3057	aaagaacagttagcatctgttggggagcagcggtcgttgggaaaagcacctctgtaacaactct	3116
Dh	3613	ccagaacgcttgctcgttggtggcagcgatggctcgttggaagagacacagatgctccagctctc	3672
Qy	3117	gcagagagacttatgaccccgctgcagaagacaaagcgtttgtagtgtgtagatgacaaga	3176
Dh	3673	ggagaggttcttatacgaacccctctggacaaggaagatgctgcttggatggcaagaataataacg	3732
Qy	3177	attgatgtataagtggtctcgcttcccaaatatagaatcgttccctcaagagccgtgtcctt	3236
Dh	3733	actgaatgttcaagtggtctccgagacacaactcgggcacatcggtcccaaggagccatctcgtt	3792
Qy	3237	caactgcagcatctgtgagacacatcgcttatgttgacaacaagccgtgtgtgctacataga	3296
Dh	3793	tgactgtcagcatctgtcagaacattgtcctatgtgagacaacagccgggtgtgtgtacaga	3852
Qy	3297	tgagatcaagaagccgcgaatgycagcaaatatccattctttatgtgaaggtctccctga	3356
Dh	3853	agagatcgttgaggccgaagaagagccacaacatatactgtcttcatcgaatgcactcctcaa	3912

[illegible]

Accession ID	Gene	Protein	Location/Qualifiers
AAZ49332	standard; CDNA; 3860 BP.		
AAZ49332			
14-MAR-2000	(first entry)		
Human wild-type multidrug resistance-1 (MDR-1) CDNA.			
Multidrug resistance; MDR-1; P-glycoprotein;			
transmembrane efflux pump; haematopoietic stem cell; transduction;			
bone marrow transplantation; chemotherapy; radiation therapy; cancer;			
gene therapy; gene replacement; genetic defect; thalassemia;			
Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;			
cytokine; wild-type; ds.			
Homo sapiens.			
Key	Location/Qualifiers		
CD5	1..3843		
/*tag=	a		
/product=	"Human wild-type MDR-1 protein"		
replace (553..555, G/T)			
/*tag=	b		
/note=	"cDNA sequence of G185V human mutant MDR-1 given in AAZ49333"		
MO9961589-A2.			
02-DEC-1999.			
27-MAY-1999;	99WO-US11825.		
28-MAY-1998;	98US-0086988.		
(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.			
Sorrentino B, Bunting K;			
WPI; 2000-072615/06.			
P-PSDB; AAY58186.			
Ex vivo expansion of hematopoietic stem cells transduced with a			
sequence encoding human multidrug resistance-1, used for bone marrow			
transplantation			



Db 1929 atccaaaagtgaattgtagcttggaaatgtcttcaaatgatccaagatccagctctaat 1988  
Oy 1882 aagagcatca-----agtcgaactctatgccaagcgctgagga 1919  
Db 1989 aagaaaagaatcaactcgttagagatgtccgtggaacacaaagccaaagaaagcttag 2048  
Oy 1920 atccccaatctaagaagataagctctccggaagctccctctatataaatttaagctt 1979  
Db 2049 taccaaaagagctcgtgtagaagataaccctccagcttcccttttggagagattagaagct 2108  
Oy 1980 aaacaagcctgaatgagccttttgggtctcgggagacattggctctgtctaaatgagac 2039  
Db 2109 aaatttaactgagatgacctatttgggttgggtatttttggccattataaatgagag 2168  
Oy 2040 tgttcaatccagatatttccatcatcttgcacaaaattataaccagtgtt---ggaaataa 2096  
Db 2169 cctgcaaccagcagatcttgcaataatatttcaaaagattataggggttttacaagaattgga 2228  
Oy 2097 tgataaaaccacattaaagcagatgagcagaattatccattccatgatatccgtcaatttggg 2156  
Db 2229 tgatccctgaacaaaacgacagaaataagtaactgttttccactattgttctagcccttgg 2288  
Oy 2157 tgttatgttctgtcagttatltcatlgaaggaattatttaacgagacagaggggaaat 2216  
Db 2289 aattattcttttattacatttttccctcaaggttttccattggaagcgtgagagat 2348  
Oy 2217 tttaacgatgagattaagaacacttggccttcaagaccatgttatccagatattggctg 2276  
Db 2349 cctccccaagcggctccgcatatcattgtttccgcatcctgacagacagagatgtagttg 2408  
Oy 2277 gtttataaaaagaaaacagacagagagcgttgacacaaatattagacatatgatatagc 2336  
Db 2409 gtttataagccttaaaaacacacactggagacattgataccagcgctccgcaatgagtcgc 2468  
Oy 2337 acaaatlcaaggagacagatgttccagatgtggcgtcttaacacaaaatgacaactaacat 2396  
Db 2469 tcaagttaaagggtctataggttccagcgctgtgtaattaccagaatatagcaaatct 2528  
Oy 2337 gggacttccagttatcatcttcccttataatagatggagagatgaatcccgatctgag 2456  
Db 2539 tgggacaggaataatatacttccatcatcgttcttgaacataacactgttactcttgc 2588  
Oy 2457 tatctccagacttgcggtgacaggaatgatacgaacccgacgaactgaactggattgc 2516  
Db 2589 aattgtaccatcatctgcaatagcagaggttgttgaatgaaatgtaagttgctcgagcaagc 2648  
Oy 2517 caacaaagataagcaagaacttaagcatgtcggaaagatagcaactggaagcttggagaa 2576  
Db 2649 actgaaagataagaagaactagaaagtgtcgggaagatcgtactgaaagcaatagaaaa 2708  
Oy 2577 tataagtactaatagtcaatttaacaaaggaagaaaccttcggacgaatgataagaagat 2636  
Db 2709 ctccggaacccgttcttcttaccaggaagcagaagtttgaacatatgttcttgcacag 2768  
Oy 2637 gcttaagactcaacagaataactcctgaaagaacacagatatgttgaagctgtatagc 2696  
Db 2769 ttgcaaggtaccatcaagaactcttgaagaaagacacacactcttgaatataatttc 2828  
Oy 2697 attcagcagccttataataatttgcctabgacagaggttctgatttggacctattt 2756  
Db 2829 cttaacccagcagatgatttcttccatgctgattgttccggttggagccactt 2888  
Oy 2757 aattcaagctgagcaatgagaccacagagggatgttcaagtcttctacgcaattgacata 2816  
Db 2889 ggttgacacataaactcaatgagcttggagatgttctgtatgatttccagctgttgcctt 2948  
Oy 2817 tggagcctatgagcagcagaaagcctgcttggctccctgataattccaagcaaatc 2876  
Db 2949 tgggtgcattgagcgttggggcagatcagttcatttgcctctgctatgccaagccaaat 3008  
Oy 2877 gggggtcgagcatctgttgccttgttggaaaagaaaccaaatatagacagccgagtlca 2936

Db 3009 atcagagccaccacalcalatgatacttgaaaaaacccctttagatgacagctacagcac 3068  
Oy 2937 aagaaggaagaaagccagacacatgtgaagggaatttagatttgcgaagctcttctt 2996  
Db 3069 ggaagagcctaattgcgcgaacacatltggaaggaatgtaacatttgttgaattgattcaa 3128  
Oy 2997 ctatccalgtccccaagatgttttccatctccgttgcgttaccctgaaglattgagcagg 3056  
Db 3129 ctatcccaaccgacccagacatccccagtgcttcaaggatcagagctcgtgaggtgaagagg 3188  
Oy 3057 aaagaacagtaacatttgtgggagcagcggctgtggggaagaaagcctctgttcaactt 3116  
Db 3189 ccaagcagctgtcctcgttgggcaagcagtgctgtgggaagagcagatgtccagctcct 3248  
Oy 3117 gcaagacttattgaccccggtgcaagagcagtgctgtttagtgttggatgtaacaga 3176  
Db 3249 ggaagcgttctacgaaccccttggcaaggaaagtgtcgttgaatgcaagaataaagcg 3308  
Oy 3177 attgaatgtaacagtggtcgtcttcccaatagcaatcgttccctcaagacctgtgctct 3236  
Db 3309 actgaatgtcagtggtcgtccgagcacacctgtgcatcgtgtccagagagccatcctgt 3368  
Oy 3237 caactgcagcatgtcgtgaacacatcgtcctatgttgaacaaagcgtgtgtgcattaga 3256  
Db 3369 tgactgcagcatgtcgtgaacacatgtcctatggaagcaaacagccgggtgtgttccagga 3428  
Oy 3297 tgagatcaagaagccgcaaatgcaagcaatatcatcttctttagaaggtctccctga 3356  
Db 3429 agagatcgttggagggaaggaagggccacacatcagcttccatcagatcagctactaa 3488  
Oy 3357 gaatatcaacacacaaagttgagctggaaggaagcagcagcttcttctggtgcgcaaaaag 3416  
Db 3489 taatatagacctaagaagtagagcaaaaggaacccagctcgtcgtgtgcgcaaaaagc 3548  
Oy 3417 actagctatgcaagggctcttcccaaaaacccaatttattgttggatgggctcac 3476  
Db 3549 catgtccataactcgttgccttgttgaagcagccctacataatttgccttggatgtaagcac 3608  
Oy 3477 ttcaagcctcgataatgacagtgaaggaaggtgttgaagatgcagatcccttgaataagcagac 3536  
Db 3609 gtcaagctctgataagaagaattgaaaggtgttgcacaaagaccccttggacaaagcagaga 3668  
Oy 3537 ggaagagcagtcgctagtgctcactacagagctcctctgcaatttcagaacgcagattgat 3596  
Db 3669 aggcgcacccgtcagattgtgattgtctcacccgctgttccacaaatccagaatgagacttaat 3728  
Oy 3597 agtgttctgcacaatgaaagataaaggaacaaaggaactatacgaagcctcctgagaaa 3656  
Db 3729 agtgggttctagaatgtgcagagtcgaaggaagcatgtgcagcatcagacagctcgtgaca 3788  
Oy 3657 tcgagacatatattttaagtttagtgaatgcacagtcag 3694  
Db 3789 gaaagcatcatttcttcaatggtcagtgctcagagctg 3826

RESULT 12  
AAN70752  
ID AAN70752 standard; cDNA: 4669 BP.  
XX  
XX AAN70752;  
AC  
XX  
XX  
DT 21-MAY-1991 (first entry)  
XX  
XX  
DE Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones  
DE Lambda HDRI0.5 and 104.  
XX  
XX Chemo-therapy resistant tumour cell; P-glycoprotein; ss.  
KN  
XX  
OS Homo sapiens.  
FH  
FH Key Location/Qualifiers  
FT CDS 425..4267  
FT /tag= a



QY	1779	ttaactgtgatgtctcccaagatatactaaagaagctfgatgaacaagatgagatgcata	1838
Db	2293	caactatgtcatcaatgacacagcaggaatgaagttgaattagaaatgaagctgtga	2352
QY	1839	ttcta-----ctgaagaagaagccaactcaactccctgcactcgtg	1881
Db	2353	atccaaagatgaattgatgcctctggaaatgctctcaatgatattcaagatccgctcaat	2412
QY	1882	aagagcatca-----agtcagacttcatctgaagaagctgagga	1919
Db	2413	aagaaaaagatcaactcgttagaggtgtccgtgatcacaaagcccaagacagaagcttag	2472
QY	1920	atccaccaatctaaagagataagttcttcctgaagtctctctatataaaatttaagtt	1979
Db	2473	taccacaagagcccttgatgaagaagtacatccctcagttccctttgagagattagaagct	2532
QY	1980	aaacaagcccggaatggcctttgtgtgtcttggggacattggctcgtgttcaaatgaaac	2039
Db	2533	aaatttaccgaatgcttattttgtgtgtgtgtattttggcattataaattgaaag	2592
QY	2040	tgatcacacagatatttcatcatcatttgcanaaataataacccatgtt--ygaataa	2096
Db	2593	ccgtgaaccacgacttgcataataatttccaagatttagagggttttacaagaattga	2652
QY	2097	tgataaaaccacacttaagaatgatagtgcagaatttatccatgatattgtcatattgg	2156
Db	2653	tgaacttgaaacaaacagaaagataagtaactgttttcaactatgttttcaagcccttg	2712
QY	2157	tgattattgtcttgcaatgatatttcatctgacagatattattttagcgaagaagggagaaat	2216
Db	2713	aatatttcttattatcaatttctcctcaaggtttcaacatttggcaaaagctgagagat	2772
QY	2217	tttaacgaatgagatlaagacacttggccttcaagcaatgttatatacagaatatgtcgt	2276
Db	2773	ccctaccaaagccgctccgatacatgttttccgactccatgctcagacagagatgtgaattg	2832
QY	2277	gtttgatgaagaagaaacacgcacgaagagcttggacaacaatatatgaacatagatagac	2336
Db	2833	gtttgatgacccctaaaaacccctcgagacattgtctacacagctcgcgaatgtgcctgc	2892
QY	2337	acaaattcaagagacaaacaggtttccagagatttggcgtcttcaacacaaaatgaaactaaat	2396
Db	2893	tcaagtttaaggggcatagatgttccagcgttgcgtlaatttaccagaatatagaaatct	2952
QY	2337	ggagccttcagttatcaatttcccttataatgataatgggaatagatcatccgatctcgag	2456
Db	2953	tgggacacgagatataatatactcccttccatctatgttggacactaaacactgttactcttagc	3012
QY	2457	tattggcccaagtatccgtcgtagcagaatgtgttbaaacccgagaacagatgttgatttgc	2516
Db	3013	aattgtatcccatcaatttgcgaatcagagagttcttgaatagaaatgltgtcttgcgaacgc	3072
QY	2517	caacaaagataaagcaagaacttlaagatgtctgnaaagataagcaactgaagctttgagaa	2576
Db	3073	actgaagaataagaagaactagaaggtctgtagaagatcgtactgaagaaatagaaaa	3132
QY	2577	tatacgttactatagtgtcatttaacaaagggaaaaagccttcgaaccaatgtataagaagat	2636
Db	3133	cttcgcgaacggttgttcttcttgactcgaagaagaagaagatttgaacatatgtatgtctcagag	3192
QY	2637	gcttcagaccccaacagaagaatccctcgagaagaacagacagattatitgaagctgttagc	2696
Db	3193	tttgcagatcatcacagaacactcttgcggaagcacacacacttttggaaattacatttcc	3252
QY	2697	attcagccatgccttataatatttgcctatgcatagcagaggttgcattttagagcctatt	2756
Db	3253	cttaccocaggcaatgtatgtatttcttccatgtctgatagtttccggttttagagcctact	3312
QY	2757	aattcaagcttgcagaaatgaccccaagaggcagtgttcatagtttttactgcgaattgcata	2816
Db	3313	ggttgccacataaactcatgagctttagaagatgttctgtagattatttcaagctgtgtcctt	3372

QY	2817	tggtgcatatgacatgagaaacgcgtcttctgtctctgtatattctcaagccaatc	2876
Db	3373	tggtgcatatgacatgagaaacgcgtcttctgtctctgtatattctcaagccaat	3432
QY	2877	gggggcctgcgcatctgttctgtctgtctgtgaaaagaaccacaatatagaacgcgac	2936
Db	3453	atcagcagcccaatcatcatctatgtcatattgaaaaaaccccttggatctgacgtacac	3492
QY	2957	agaaagcaaaaagccagacacatgtgaagggaaattagaatttcgagaagctctctt	2996
Db	3493	ggaagcctaatgtccgaacacatctggaagaagaatgtccacatttggatgaaatttcaa	3552
QY	2997	ctatccatgttgcgccaaatgtttcaatccctgcgtgtgtatattccctcagtattgagcag	3056
Db	3553	ctatcccaaccgcagccagacatcccaatgtgtcttaaggaactcgagcctgagagtgaaagag	3612
QY	3057	aaagacagtatcatcttgttgaggagcgcgcgtctggaagaacacatctgttcaactct	3116
Db	3613	ccagaacgtctgtctctgttgcgcgcgcgtgtgcgtgtggaagagacacagtggtccagctct	3672
QY	3117	gcagagacttataagaccgcgtgcgaagacaaatgtctgtttgaatgggtgtgaaatgaaga	3176
Db	3673	ggagcgttcttaagcaccctcttgcaagggaagtgctgtgttgatgcaagaataaagcg	3732
QY	3177	attgaatgatacgtgtgtctccgttcccaataagcaatcgtctctcaagagcctgtgtctt	3236
Db	3733	actgaatttcaatggtgtctccgaacacacgcgtggtcactgtgtccagagagccatctctt	3792
QY	3237	caactgcagcacttctgaagacatctgcctatgtgtgaacacagcgtgtgtgttcattga	3296
Db	3793	tgactgcagcacttctgaagacatctgcctatgtgtgaacacagcgtgtgtgttcattga	3852
QY	3297	tgagatacaagaagccgcgaatgtgcagaaatatactatcttatttgaaggtctctctga	3356
Db	3853	agagatcgttgaggcgagcaagaagagccaaacatactgcttcaatcagatcactgtctaa	3912
QY	3357	gaataacacacacaatgtgtacttgaaagagacacagcttcttgccgcccgaagaacaag	3416
Db	3913	taaaatagtcactaaagtlagaagacaaaggaactcagctctctgtgtgcccagaacaacg	3972
QY	3417	actagctatccaagggcctctctccaaaaccacaatcttatgttgaatgaagccac	3476
Db	3973	catgtccatagctctgtgctctgtttgaacacgcctcatatttgccttggatgaagccac	4032
QY	3477	ttcagccctcgataatgacsgtgaagagtggttcagcatcgtcccttgataaagccagac	3536
Db	4033	gtccagctctgatacagaagaagtgaaaaggtgttcacaagaagcccttgacaagaaccagaga	4092
QY	3537	gggaagacacagctctatgttcaactcaacagagctcttgcaattgaaagcgcagattgat	3596
Db	4093	aggccgcacctgcgatgtgtatgtgtccgcctctgcaccatccagaatgcagactaat	4152
QY	3597	agtgttctgacatctgaaagataaaggaacaggaactcatcagaagctctctgaagaa	3656
Db	4153	agtgttctgacatctgaaagataaaggaacaggaactcatcagaagctctctgaagaa	4212
QY	3657	tcgaagacatatatttaagtatgaatgaatgcacatcag	3694
Db	4213	gaaagcatatttcttcaatgtgcagatggtccagagctg	4250
RESULT	13		
AAZ24042			
ID	AAZ24042	standard; cDNA; 8630 BP.	
XX	AAZ24042;		
AC			
XX			
DT	04-FEB-2000	(first entry)	
DE			
XX			
XX			
KW		Retroviral vector; gag gene; gene therapy; chemotherapeutic agent; hematopoietic stem cell transformation; mdr-1; ss.	



Db	3028	tggagtcattgttggagaaagaaatcagatgaactcatgaagagaaggcatttctt	3087
QY	1779	ltcaacttgtatgtcaccagagatatataaaagctgtatgaacagatgtaatagtacata	1838
Db	3088	caaaactgtcacatgcagacagcaggaatgaagtattgaataatgcagctgtatga	3147
QY	1839	ttcta-----ctgaagaagaaccactcaactcttccctgcgactcgtg	1881
Db	3148	atccaaagaatgaatgtatgccttggaaatgtcttccaatgtatcaaatgcagctcat	3207
QY	1882	aagagcatc-----agtcagactcatcttgcgaagctgtggga	1919
Db	3208	aagaaagaatcaactcgttgtaggtccgtgtacacaaagcccaagacagaagcctta	3267
QY	1920	atccacccaatctaaagagataaagttcctctgaagtcctctatbaaaattttaagt	1979
Db	3268	taccaaagagctcctgtatgaagtaatactccacagttcccttcttgtaggtatgaagct	3327
QY	1980	aaacaagccttgaatgccttttggttcttgggacatcttgctcttctaattggac	2039
Db	3328	aaatttaactgaagtgagccttatttctgttgggtatatttggccattaaatgagag	3387
QY	2040	tggtcatccagatatttccatctcttgcataaataatcaaacagttt---ggaaataa	2096
Db	3388	ccgtgacaccgcgatttgcataaatatttccaagatataaggggttttaacaagaattga	3447
QY	2097	tgataaaacacattaaagcagatgacagaatattatccatgatatctgcatttggg	2156
Db	3448	tgatccctgaacaaacagacagatagtaactgttttcacattgttttcagcccttgg	3507
QY	2157	tgattattgtcttgcagttatttcatlctgaaggatattttacggcgaagcagggaat	2216
Db	3508	aattattccttcttcatcttcttccatcaaggttccacatttggcaagcttggagagat	3567
QY	2217	tttaagatgagatgaagacacttggccttccaagcagttatatcggatattgacctg	2276
Db	3568	ccctcaacagcgctccgataacttggtttccgatccatgcctcagaacgagatgtgagttg	3627
QY	2277	gtttgatgataaaggaaaaacagcacagagagctgtacaacaatatatgacatagatatagc	2336
Db	3628	gtttgatgaccccttaaaaaacacacacttggagcatgtactacagctgcgaatgatgtgc	3687
QY	2337	accaatltcaaggagcaacaggttccagagattggcgtcttaacacaaatgtcaactaacat	2386
Db	3688	lcaagttlaaagggtctataggtcttcacagcttgcgtgaatttaccacagaatatagacaactc	3747
QY	2397	gggacttccagtatcatcttcccttatataatgatgagtgaggagaaactcccatctctgag	2456
Db	3748	tgggacacaggaataattatcatccttcatcctatcgtgttggcaactaaacactgtaacttcagc	3807
QY	2457	tattgtccagttacttgcgtgtacaggaatgatgtgaacccggacgaatgactgatatgttc	2516
Db	3808	aattgtaccatcatctgcaatagcagaggtgtgtgaatgaaatggttgcctggacaagc	3867
QY	2517	caacaagaataagcaagaacttaagcatgctgtgaaagaatagcaactgaaagctttggagaa	2576
Db	3868	actgaagaataagaagaactaagaaggtctgtggaaagatcgcgtactggaagcacaataagaa	3927
QY	2577	tataagttacatagttgtcatttaacaaggagaaagccttcggcgaacatgtatgaagaat	2636
Db	3928	cttcggaacccgttgttcttcttgcgtcaggaagcaggaagtttgaacatagtatagctccagag	3987
QY	2637	gcttcaagactcaaacagagaataacctcogaagaagacacagatattatgaagctgttatgc	2696
Db	3988	tttgaaggttaccatacacagaacaccttttggagaagaacacacatccttggaaattaacttttc	4047
QY	2697	attcagccaagccttatataatttgcctatgtcagcagggttgcgatttggagccattt	2756
Db	4048	cttccaccagcgaatgattatatttccatagctgtgattgttccggttttggagccctactt	4107
QY	2757	aattcaagcttggacgaatgacccacagaggcagtcttcatagtttttactgcgaattgata	2816
Db	4108	gtttggacaataaactcttgagcttggaggatgtcttcgtttatgtatttccagcgttgcttt	4167

QY	2817	tgagcatatgccaatgcgaaaaacgcctcgttttgctctgtatattccaaagccaatc	2876
Db	4168	tggtgcatacggccgtggggaaagtcattcatttgcctctgtcattgccaagccaaat	4227
QY	2877	gggggcctgcgactgtgttcctctgtcttgaaaaagaaacaaatatagacacgcgac	2936
Db	4228	atcagcagcccaatcatcatatgatcatgtgaaaaaaccccttggatlgcagctaacac	4287
QY	2937	agaagggaaaaagccagacacatcgtgaaagaaattagaattcgagaagctcttcct	2996
Db	4288	ggaagccctaagtcgcgaacacatgtgaagaaatgtccatttggtagaattgtattcaa	4347
QY	2997	ctatccatgtgcgcccaatgttttcaactccgttggtctatccctccagtatgtgaacgag	3056
Db	4348	ctatcccaaccgcagccgaacatcccaagtctccagggaaccggaacccggaagtgaagaag	4407
QY	3057	aaagcagatcaccttcttgaggagcagcgcgtctgaggaaaagacattcgttcaactctc	3116
Db	4408	ccagacgcgttcctctgttgaggcagtgccgttgaggaaagacagtgcttcagctcct	4467
QY	3117	gcagagacttatgaccccgctgcgaaggaacaaatgctgttttgatgtgtgtgatagc	3176
Db	4468	ggaagcggtcttaccgaacctcttgccaggaagaaatgctcgtcttgatggcaagaataaagcg	4527
QY	3177	attgaattgaagtggtctcgcttcccaaatagacaatcgtcttccacagaacgcgttgctct	3236
Db	4528	actgaatttcaatggtctccgagacacacctgggcacacgttccacggagcccatctgt	4587
QY	3237	caactgcagcatctctgagagacatcgcctctatgtgacaaacagccgtgtgtgtca	3296
Db	4588	tgactcagcacttgcagagacattgccttgagacaaacagccggtgtgttcatcaga	4647
QY	3297	tgagttcaaaagaacccgaatgcagaaatatccacttctttatgtgaaggtccctga	3356
Db	4648	agaagctctgagggcagccaagaaggccaaacatacatcgtcttcatcagatcagctctaa	4707
QY	3357	gaatatacaacaacaanaatttgcagctgaaagagacacagcttcttcgctgcagaaacaa	3416
Db	4708	tcaatatatgccttaagtagagacaaagaaactcaactgcctctcgtgtgscagaaacaa	4767
QY	3417	actagctattgcgaaggcctcttccaaaaacccaatttatgttgcatagtgagccac	3476
Db	4768	cattgcctaagctcgtgcctctgtttagacacagctctcatatttggatttgcatagagccac	4827
QY	3477	ttcagccctgcataatgacagtggaaggtgtgtcagcgtccctctgtatataagccagac	3536
Db	4828	gtcagctctgcgatacagaagaatgaaaagttgtccagaagccctgcgaacaaagccagaga	4887
QY	3537	gggaagagcagcctagtggtcattccacaaagctctctgcgaattcagaacgcagattgat	3596
Db	4888	aggccgcacactgattgtgattgtctaacgcgtcttcacatccagaatgcagacttaat	4947
QY	3597	agtggtctgcacaatgaaagataaagaaagaaagaaatctatcacaagatcctcttga	3656
Db	4948	agtggtgttctcaagatgtgcagagtcgaagacagctatggaacgcacagcagctgtgcaca	5007
QY	3657	tcgagacatatatttaagtagtgaaatgcacagctcag	3694
Db	5008	gaaagagcatcatatttccaatgtgtcagtgctccagctgt	5045
RESULT 14			
AAZ49333			
ID	AAZ49333 standard; cDNA; 3860 BP.		
XX	AAZ49333;		
XX	AC		
XX	DT		
XX	14-MAR-2000 (first entry)		
DE	Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.		
XX	Multidrug resistance; MDR-1; P-glycoprotein;		

RESULT	14
AAZ49333	
ID	AAZ49333 standard; cDNA; 3860 BP.
XX	
AC	AAZ49333;
XX	
DT	14-MAR-2000 (first entry)
XX	
DE	Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX	
KW	Multidrug resistance; MDR-1; P-glycoprotein;



XX	transmembrane efflux pump; haematopoietic stem cell; transduction;
KW	bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW	gene therapy; gene replacement; genetic defect; thalassemia;
KW	Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW	cytokine; mutant; ds.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	1..3843
FT	/tag- a
FT	/product- "Human G185V mutant MDR-1 protein"
PN	W09961589-A2.
XX	
XX	02-DEC-1999.
PD	
XX	
XX	27-MAY-1999; 99WO-US11825.
PE	
XX	
PR	28-MAY-1998; 98US-0086988.
XX	
XX	
PA	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX	
PI	Sorrentino B, Bunting K;
XX	
DR	WPI; 2000-072615/06.
DR	P-PSDB; AAY58187.
XX	
XX	
PT	Ex vivo expansion of hematopoietic stem cells transduced with a
PT	sequence encoding human multidrug resistance-1, used for bone marrow
PT	transplantation -
XX	
XX	
XX	Example 1; Page 79-82; 113pp; English.

PS Example.; Page 79-82; 113pp; English.

XX

CC This sequence represents cDNA encoding human G185V mutant multidrug  
CC resistance protein MDR-1, where the Gly residue at position 185  
CC of the wild-type protein (AAV56186) is replaced by Val. MDR-1 is a  
CC transmembrane efflux pump, responsible for the export of drugs from  
CC cells, particularly cancer cells. The wild-type MDR-1 shows increased  
CC resistance to etoposide and decreased resistance to vinca alkaloids  
CC compared with the G185V mutant. The invention relates to transducing  
CC haematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified haematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC haematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from haematopoietic stem cells, e.g., thalassemia,  
CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in haematopoietic stem cells.  
CC Haematopoietic stem cells that express MDR-1 will be protected against  
CC chemotherapeutic agents, so can be engrafted while the patient is  
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells  
CC provides sufficient cells to permit standard biochemical analysis.  
CC Overexpression of MDR-1 allows cytokine-driven expansion of  
CC haematopoietic stem cells by at least 10-fold compared with a maximum  
CC of 4-fold in known procedures.

XQ

Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Query Match	33.08	Score 1222	DB 21	Length 3860
Best Local Similarly	60.88	Pred. No. 0		
Matches 2176	Conservative	0	Mismatches 1270	Indels 132
				Gaps 6

Oy 162 caggttgacctgataatglttggaaagtggtgtgcgcctgatttggtaataaa 221  
 Db 336 cagataagccataatcaacagtcgaacttgcgcggagtgctgctgcgcttaacatca 395  
 Oy 222 gatttccttgtagataaactgagcagcagacccaagagatttcgaataacagittt 281  
 Db 366 gattcaatttggcgctgcgcgtgcgaagacaataacacaaatttgaataacagittt 455

QY	282	tcattcaagtttttgcgaaggaacatcggtcggtttgtatagctgtgacatcggttgaaactaa	341
Db	456	tcatgtactataatgtgcgaagaagatagagctgtgtttgtgtgtgcagatgtctggggagactaa	515
QY	342	cactgcgcatacgaacacatgtac---aaatlcagatgtgtatcttggaagataaagatgtcct	398
Db	516	caccgcgacttacagatgtgtctcttaagatataatgtatattgtgtgcacaaatattgat	575
QY	359	gttgtttccaacaacatgtcctactcttttcogatttgcgtcgcgagttgtttgtgtgaaggtctg	458
Db	576	gtctcttcaagcaatgtgcacaacattttccactctggttttaagatagatttaacacgtgtgtg	635
QY	459	gaaactccccttaagtgtcctatccacagtcctccctcttaattgaatgtcctcgcgcagctg	518
Db	636	gaagctaacccctgtgtgattttgtgcacacgtccctgttctcttgtaactgtcgaactgtcgtctg	695
QY	519	tctctagaatgtgtcatccatcttgacccagtaaggaattaaatgtccatltcccaaaagctgtggc	578
Db	656	ggcaaaagtactatcttcattcaatctgaataaagaacctcttagctgtcaaaagctgtggc	755
QY	579	tgtgtgcagaagaagttctgttgcatacgcgaagtcacatgaccttaagccttaaggcccgagaa	638
Db	756	agtagctgaagaaggtctctgtgcagcaaatctgaagacgtgtgacttctgtgtgagagcaaaaga	815
QY	639	agaacttcaaaagtccttccctttaataataacaagatatgtctgtttattattcccca	698
Db	816	aggaac-----	820
QY	699	gttgcgtactaagtgtgtctgttntttgttaagatatacagaagactccaagatgtcga	758
Db	821	-----ttgaagaagtcacaacaaatattgaagaagctaa	854
QY	759	ggattttgtgcataaagaagactatagttcttaaaagtgctctctgtgtgtgtgtactctct	818
Db	855	aagaaattgtgcataaagaagactatacgaagcaatatcttcataagtggtgcgtcttccgtc	914
QY	819	tatgatattgaaccttatgtgactgtctttcttgta tggaaacctctgtattcttaatgtgaga	878
Db	915	gactcatgtcatcttatgtcctgtgcctctctgtatgtggaaacctgtgtcctctcagggga	974
QY	879	acctgtgatatccatcggtgactgtctctgtgtcttctctttagtgttaatccataagcaagta	938
Db	975	-----atattcttatgtgacaagtaaccatccatctatcttcttcttgtaattatggggctt	1028
QY	939	tgtgcattgtgagcaagaagctccctccactcttgaaaccttcgcataagcccgagagctgcct	998
Db	1029	tagtgttttgacaagcatctcccaagcatgtgaagcatcttgcaaatgcaagagggagcgctta	1088
QY	999	tcaatttttccaggttattgtataaagaaccagatataagataaactttccacagctgagata	1058
Db	1089	tgaactcttccaagataattgtataaagccaagatattgacaogactattcgaagaagtgaggca	1148
QY	1059	ttaacctgaatcccatagaaaggaactgtggagatttaaaaagtcttcttaattatccatc	1118
Db	1149	caaaacagataatataatgaagggaacttgggaattccgaagaagtccacttccacttaagtlaccac	1208
QY	1119	aagaccacatctacaagattcttguaaagttccgaaatctcogaagataagttctgtgagagacagt	1178
Db	1209	tcgaaagaagattgaagatcttggaaaggccttgaacctgaaggtgtgcagagtggggcagagcgt	1268
QY	1179	cgctctgtgtgttctcaattgtgcagtgtggaaagatagcgtatagttcccaacttctgcagaagtt	1238
Db	1269	ggcccttggttggaaacagtgtcgtgtgggaagagcaacaacagttccagctgtatgcagaagcct	1328
QY	1239	atattatcccggaatgtatgtgtcttaacagtggtgagatcgaagatgaatacagaagctttaaagt	1298
Db	1329	catatgaccacagaggggagatgtgcagtgtrgttgtgcacagatataatgagacataaatgt	1388
QY	1299	gcggcatcatatcgagaaccaataattgtagtgtgtatgtaacagagccgttttgtctgggacac	1358
Db	1389	aaggtttcttaacgggaataacattgtgtgtggtgagatcagaagacctgtattgtttgtccacac	1448

QY	1339	catcgatlaacaatatcaagtagtagcgacagatgatgtgactgtatgtagaagatltgagagac	1418
Db	1449	gatactgtgaacaacatctgcgtatgcccgtlbaaatagtccacatgtgatltgagaac	1508
QY	1419	agcaaggagagcaaatgctgatgtatttgcatagggttccctaataattatcatc	1478
Db	1509	tgtcaaggagacccaatgacctatgactttatcaagaacctgcctccaataatttgcacct	1568
QY	1479	ggtaggggaaaaagagagctcaaatgtgtgagggcgagaaacagagagatcgcgaatttcg	1538
Db	1569	ggttggagagagaggggccagttggtgtgtggcgaaagcagagatcgcgaatttcgacg	1628
QY	1539	tgccttagttcgaaaccccaagatttcgatlttagatagggtagctgtgccttgatlc	1598
Db	1629	tgcctgtgttcgcacaacccaagaacctcctcctgtgtatgtagccacgtlcaagcctttgacac	1688
QY	1599	agaaagcaagtagcgttcaagctgcactgtggaaaggcgagaaaggttgcagctcaat	1658
Db	1689	agaaagcgaagcagtggtltcaagtggtcctcgtgataaggccaagaaggtcggacacacat	1748
QY	1659	cgltgtagacacccgaacttctactatctcgaagtgtgacatttattgttgaccctaagga	1718
Db	1749	tgttagagctcatcgtttgtctacagttcgttaatgtgcagtcatacgtcgtgttgatga	1808
QY	1719	tggaaatgtcgtcgagagaagagacatgtcgaactaattgagcgaacagagttcatatta	1778
Db	1809	tggagatcatgtgtgagaagaagaaatcatgtatgaacctatbaaagaaagagcatlactt	1868
QY	1779	ttaactgtgattgtcacaagatattaanaagctgtgtgaacagatgtgagtcgaatgcata	1838
Db	1869	caaactgttcacaacatgcagacagcaggaataatgaagttgaattgaanaatgcagctgatga	1928
QY	1839	ttcta-----ctgaagagaagacaactcaacttccctcgtcacctgtg	1881
Db	1929	atccaaagtgaaattgatgtccttggaaatgtcttcaaatgtatcaagatccagctcaat	1988
QY	1882	aagagcatca-----agtaagacttcatgtgacaaggtcgtgga	1919
Db	1989	aagaaaaaatcatcaactctgtgagatgtlccgtgtgataccaagcccaagacagaagaagcttag	2048
QY	1920	atccaccacatcctaaagagataagtcctcctcgtgaagctccctctatbaaattttaagtt	1979
Db	2049	taccaaagaagcctctgagatgaagaatataccccaagttcccttltggaggaattagaagct	2108
QY	1980	aaacaagcctgaatggccttlttggttctcggggagacatlgctctgttcttcaatgtgaac	2039
Db	2109	aaatttaactgaatgacctattttgttgtgtatatttggccattabaatgtgaggg	2168
QY	2040	tgttcatccaagtatttccaatcacttctgcaaaaataataacaatgtt---ggaataa	2096
Db	2169	cctgcgaaccaagcatttgcataataatttccaagaattatagggtttlttcaagaattga	2228
QY	2097	tgttaaacacacatthaagaatgagtgcagaatttatttccatgatatttgcataattggg	2156
Db	2229	tgaatcctgaaacaaacgacagatatagtaactgttttccaactatgttttccaagccttgg	2288
QY	2157	tgtatttgccttgcacgttatatttcatgtccagggatatttcaagcgagagcagggagaat	2216
Db	2289	aattatttcttatttaacatttttccctcaagtttccaacttggcaagctgtgagagat	2348
QY	2217	tttaacgatgatattaagacacttgcctccaagaagccatgtttatatacaagatalgtccgt	2276
Db	2349	cctcaccaagcggctccgatacattgtttccagttcatgtcgcagacagatgtgagttg	2408
QY	2277	gtttgtatgaaaagaaaacacgcacagagagctgtgacaacaatatattagccaatagatagc	2336
Db	2409	gtttgtatgacctaataaacacacactgtgagatgttacttaccaggtcgcgaatagtgtgc	2468
QY	2337	acaattccaaggagcaacaggttccagagattgtggctttaaacaanaatgaactaatcat	2396
Db	2469	tcaagtttaaagggtctatagtttccaggtctgtctgtatattaccagaatatagcaaatct	2528
QY	2397	ggagacttccagttacatcttccattatatabtgtatgggaatgacatctccgtacttcgag	2456

Db	2529	tgggacaggaatataatataatcccttaacatctatggttggcaactaacactggttactcttagc	2528
Qy	2457	tattgcccacgaactactggtccgttgacaggaatgatatgaaaccgacgaatgacgtgatttgc	2516
Db	2589	aattgtatccatcatattgcaatagcaggttgttgtaaatgtaaatgtgtctgtgacaagc	2648
Qy	2517	caacaaagataagcaagaacttaagcatgctgtggaagaatagatgaactgtgaagctttggaa	2576
Db	2649	actgaagaataagaagaactagaaggttgcgtggaagatgcgtactactgaagaacataga	2708
Qy	2577	tatacgtactatggtgttcattaaagggaaggaacgttcgtgacgaatgatatgaaagat	2636
Db	2709	cttcgcgaacggttgttcttcttgactgaagagcagaagtttgaaactaagtatgctcagag	2768
Qy	2637	gtctcgaactcaacaacgaataatccctcgagaagaacacagaattatgttgaagcttlatgc	2696
Db	2769	tttgcaggttatccataacagaaactcttggggaagcacaacatcttggatattacatttc	2828
Qy	2657	atlcagccatgcttatataattttgactatgacgaggtttgcatttggagctctatc	2756
Db	2839	cttcacccgaagcaatgattgtattttctcctatgctgattgttccggttttggagcctact	2888
Qy	2757	aattcaagcttggagcaatgaccccaaggcgatgttcaatgatttactgaatttgata	2816
Db	2889	gggtggacataaactcctatgagctttgagagatgtccgttaagatttcaatcagctgttgcct	2948
Qy	2817	ttggagcatatggtccatcgcgaaanaacgctcgttlttgctcctcgatataatlcacaagccaatc	2876
Db	2949	ttgtggccatgtgcgttggggaagtgcagttcaattgtcctctgactatgccaagccaat	3008
Qy	2877	gggggctgtgcatactgtttgtccttgttggaaaagaacccaatatagacagccgagta	2936
Db	3009	atccgagcccaatcatctcatatgatacttgaaaaaccccttggatttgacgtcatacagac	3068
Qy	2937	agaagggaaaaagccacacacatgtgaagggaatttaggtttcgaagaagctctctct	2996
Db	3069	ggaaggtgctaatgctcggaacacatltggaaggaatgtcacatttggvtgagttgtatcca	3128
Qy	2997	ctatccatgtgcacacatgatttttcatccctgcgttgcctatccccaatlgtggcgag	3056
Db	3129	ctatcccacccgacgcgaactcccagttgtctcagggactgagcctcgagggtggaagag	3188
Qy	3057	aaagacagtagcatttgttggggagcgagcggttgggaaagacactctgttcaacttc	3116
Db	3189	ccagcgctgtgctctgttgggcagcggtgtgttgggaagcagacagctgtgccactctc	3248
Qy	3117	gcagagacttataatccccgttgcagaagacaaatgtcgttttggatgtgtgataatga	3176
Db	3249	ggaggtgttctactacgaccccttgcagaggaagatgtcgttcttggcaagaataaagcg	3308
Qy	3177	attgaatttaagtggtctgccttcccaatagaatctgttccctcgaagcgctgtgctt	3236
Db	3309	actgaatttcaatgtggtcctccgagacacactgtggcatcgtgtcccgagggcccatctgtt	3368
Qy	3237	caactgcagcatlttgtagagacatcgtcctatgttggacaacagcgltgtgtgcattaga	3296
Db	3369	tgactgcagcatltgtgagaacttgcctatggagacacagccgggtgtgtgtcaagga	3428
Qy	3297	tgaatcaagaagccgcaaatgcagcaatatccattctttatitgaagtcttccctga	3356
Db	3429	agagatcgttggggcgacgaaggaaggtccacatatactgtcctcatcagatcctgtccaa	3488
Qy	3357	gaataacacacacaaatgttgactcgaaaggaagacagcttcttgcgcgcgaanaacag	3416
Db	3489	taaatatagcactaaatagtagagacaagaatccacagctctctcgtgtgtgcgaacaacg	3548
Qy	3417	actagctattcaagggctcttcccaaaaacccaatlttatgttggatgagggcac	3476
Db	3549	catgtccatagctgtgctcttgttggacagcctcatatttgccttggatgaagggcac	3608
Qy	3477	ttcagcccttcgataatgacgttgaagaagtgttctagcatgccttgataaagccagga	3536





```

Db 4828 gtcagctctgtgatacagaagtgaaaagggttgccaagaagccctggaacaaagccagaga 4887
      ||||| || ||||| || ||||| ||||| || ||||| || ||||| ||
QY 3537 gggaaggacatgctctagtggtcactcacagcctctctgaaltcagaacgagattgat 3596
      || || || || || || || || || || || || || || || || || || || || ||
Db 4888 agccgcacacctgcatlgtgaltgctcaccgctgccaaccatccagaatgcagacttaac 4947
      || || || || || || || || || || || || || || || || || || || || ||
QY 3597 agtggctctgcacaatggaagataaaggaaagaaatcctcaagagctctctgagaaa 3656
      || || || || || || || || || || || || || || || || || || || || ||
Db 4948 agtgggtlctcagaatggcagagltcaaggagcatlggcagcatcagcagctgtgtgcaca 5007
      || || || || || || || || || || || || || || || || || || || || ||
QY 3657 tcgaacatatattttaagttagttgacatgcagtcag 3694
      || || || || || || || || || || || || || || || || || || || || ||
Db 5008 gaaagcatctatttttcaatggtcagtgctcaggtg 5045
      || || || || || || || || || || || || || || || || || || || || ||

```

Search completed: April 22, 2002, 23:07:44  
 Job time: 26052 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:48:25 ; Search time 14771.4 Seconds

(without alignments)  
4134.511 Million cell updates/sec

Title: US-09-873-409-13

Perfect score: 3702  
Sequence: 1 ttccgcttgcgcagtgact.....atgcacagtcagtcagtcga 3702

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_ov:\*

5: gb\_pat:\*

6: gb\_ph:\*

7: gb\_pl:\*

8: gb\_pr:\*

9: gb\_ro:\*

10: gb\_sts:\*

11: gb\_sy:\*

12: gb\_un:\*

13: gb\_vl:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_om:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_sy:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htgo\_hum:\*

31: em\_htgo\_inv:\*

32: em\_htgo\_rod:\*

33: em\_htg\_hum:\*

34: em\_htg\_inv:\*

35: em\_htg\_rod:\*

36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1234.4	33.3	4279	6	AX105078	AX105078 Sequence
2	1234.4	33.3	4279	6	AX105080	AX105080 Sequence
3	1234.4	33.3	4279	6	AX105082	AX105082 Sequence
4	1232.8	33.3	4279	6	AX105057	AX105057 Sequence
5	1231.2	33.3	4045	12	AF269224	AF269224 Synthetic
6	1229.6	33.2	4317	4	AF045016	AF045016 Canis fam
7	1229.6	33.2	4317	6	AX105059	AX105059 Sequence
8	1225.2	33.1	4646	6	I49610	I49610 Sequence 2
9	1225.2	33.1	4646	9	HUMMDR1	M14758 Homo sapien
10	1225.2	33.1	4669	6	I08557	I08557 Sequence 3
11	1225.2	33.1	6505	6	AR028671	AR028671 Sequence
12	1225.2	33.1	9318	6	AR028672	AR028672 Sequence
13	1223.6	33.1	4378	6	E02326	E02326 Multidrug r
14	1223.6	33.1	8630	6	AX012321	AX012321 Sequence
15	1220.4	33.0	8630	6	AX012320	AX012320 Sequence
16	1217.2	32.9	4669	6	AR055785	AR055785 Sequence
17	1217.2	32.9	4669	6	AR091275	AR091275 Sequence
18	1215.6	32.8	4186	6	AX108654	AX108654 Sequence
19	1215.6	32.8	4195	6	AX108656	AX108656 Sequence
20	1214.2	32.8	4264	6	AR051647	AR051647 Sequence
21	1212.6	32.8	4264	6	AR051650	AR051650 Sequence
22	1209.2	32.7	3988	6	AX024454	AX024454 Sequence
23	1209.2	32.7	4192	9	AF016535	AF016535 Homo sapi
24	1189.2	32.1	4018	5	GG89799	GG89799 Gallus ga
25	1185.2	32.0	3924	6	AX024455	AX024455 Sequence
26	1185.2	32.0	4296	10	HUMMDR3	M23234 Human membr
27	1185.2	32.0	4296	10	CRUGP1	M60040 C.gri-seus p
28	1183.6	32.0	3987	10	CRUGP1185	M59254 Chinese ham
29	1183.6	32.0	4304	10	CRUGP1165	M59253 Chinese ham
30	1181.8	31.9	4280	10	CRUGP11	M60041 C.gri-seus p
31	1162.8	31.4	3858	4	OA078609	U78609 Ovis aries
32	1127.6	30.5	4084	10	MUSMDRA	J03398 Mouse mdr g
33	1127.4	30.5	4927	10	AF257746	AF257746 Rattus no
34	1125.4	30.4	4298	10	MUSMDR	M14757 Mouse multl
35	1124.8	30.4	4323	10	AF286167	AF286167 Rattus no
36	1121	30.3	4233	6	AR123273	AR123273 Sequence
37	1118	30.2	3905	10	CRUGP111	M60042 C.gri-seus p
38	1113.2	30.1	4924	10	MUSMDRIA	M33581 Mouse p-gly
39	1111.8	30.0	3489	4	AB029153	AB029153 Felis cat
40	1111.6	30.0	4356	10	MUSMDRA	M30697 Mouse multl
41	1093.8	29.5	4254	10	RATMDRM	M81855 Rat mdr mRN
42	1093.6	29.5	4390	5	XLU17608	U17608 Xenopus lae
43	1089	29.4	3912	10	RATPGLYCO	L15079 Rat p-glyco
44	1048.4	28.3	3682	10	MUSPGLYIA	M24417 Mouse phosph
45	856	23.1	2726	6	I33621	I33621 Sequence 1

## ALIGNMENTS

RESULT	1	
LOCUS	AX105078	
DEFINITION	Sequence 22 from Patent WO0123540.	PAT
ACCESSION	AX105078	
VERSION	AX105078.1	GI:13921228
KEYWORDS		
SOURCE	dog.	
ORGANISM	Canis familiaris	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.	
AUTHORS	Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.	
TITLE	p-glycoproteins and uses thereof	
JOURNAL	Patent: WO 0123540-A 22 05-Apr-2001;	
FEATURES	LOCATION/Qualifiers	
source	1..4279	
	/organism="Canis familiaris"	





Db 1891 CAACTTGTCAATGACAGACAGAAATGAATTGATTAGAAATGGCACGCTGCA 1950  
 Qy 1898 attctactgaaagaagcaactcaactctctcgtcactcgtgaaag-----cat 1889  
 Db 1951 ATCCAAAGGTAAAGTGGCTTGGAAATGCTCCAAAGATCCAGGCTCAGTTTAAT 2010  
 Qy 1890 caagtcagactctatgcaagcgtgagaaatcccaactcaaaagataagctc-- 1947  
 Db 2011 AAAAAGAGATCAACTCCGACGAGTATACATGCACCAAGGCCAGACAGAAAGCTTGG 2070  
 Qy 1948 -----cctgaagctcctatataaatttaaaagt 1979  
 Db 2071 TACAAAGAGACCTTGATGAGATGTACCTCCAGTTCCCTTCGAGAGATTTGCAACT 2130  
 Qy 1980 aaagaagccgtaatgagccttctggtcgtgagacatggtcctcgtcctaalgagac 2039  
 Db 2131 GAACCTCACTGAATGGCTTATTTGTTGGTATTTTGGCTATTTTAAACGGAGG 2190  
 Qy 2040 tgtcatccagtatcttccatcatccttgcgaataataataccatgctt---ggaataa 2096  
 Db 2191 CCTGCAACACACATTTTCAATTAATTTTCAAGATTAAGGATCTTTACCCGAGATGA 2250  
 Qy 2097 tgataaaacccatlaaagcaatgacagaatattatccatgatatcgtcaatttgg 2156  
 Db 2251 GGATCTCGTAACCAAAACACAGATAGTACATGTTTCTGTATTGTTCTAGCTTGG 2310  
 Qy 2157 tgtatttgccttgcagtaattcaatgaggaattttaaagcagagagggaaat 2216  
 Db 2311 AATTAATTTCTTTTATTAATTTTCTCCAGGCTTCAATTTGGCAAAAGGGGAGAT 2370  
 Qy 2217 tttaagctgagattaagacacttgccttcaagcagatgatatcagagatatgctg 2276  
 Db 2371 CCTCACTAGCGGCTTCATATGATGTTTCAGATCCATGCTGAGACAGAGATGAGCTG 2430  
 Qy 2277 gtttgaatgaaaggaaacacacagagagccttgaacaacaatlaagcatagatatag 2336  
 Db 2431 GTTTGATGACCTTAAACACACAGATGAGATGACAAACAGGCTTGGCAATGATGCGGC 2490  
 Qy 2337 acaaatcaagagagacagatccagagatggtgccttcaacaacaatgcaactaacat 2396  
 Db 2491 TCAAGTTAAAGGGGCTATAGGTCCAGGCTTCTGCTATTAATGCAATATGCAAACT 2550  
 Qy 2397 gggacttcaattatcattccttataatagatgagagatgacatccgattcgtag 2456  
 Db 2551 TGGGACAGGCATTAATATATCTTAATCTATGTTGGCAATTAACCTTACTTATGAC 2610  
 Qy 2457 tattgtccagtaactgctgcagagagatgataaagcgaagcaatgagactgttgc 2516  
 Db 2611 AATGTACCCATTCATTCGAAATAGCAGAGTGTCAAAATGAAATGTTGCTGACAAAGC 2670  
 Qy 2517 caacaagaataagagaacttaagcaltgctgaaagaatagaactgaaagcttggaa 2576  
 Db 2671 ACTGAAGATTAAGAAAGAGCTAGAAAGAGCTGGGAAATGCTTACAGAAAGCATGSA 2730  
 Qy 2577 tatagctactatgctgataaagaagaaagccttcagcgaatgataagagat 2636  
 Db 2731 CTTCGGAAGCTGTTGTTTGTGACCTGGGACAGAAATTAATACATGATGACACAG 2790  
 Qy 2637 gcttcaagctcaacaagaataatcctgaagaagaacagatatgtgaagctgtatgc 2696  
 Db 2791 TTTGCAATACATACAGAAACTCTTTAGAGAAAGCACACATCTTGGGCTCTATTTC 2850  
 Qy 2657 altcaagcaltgcttataatcttgcctatgcagcagaggttgcatttggagcattc 2756  
 Db 2851 TATCACCCAGGCATATGATGATTTTCTATGCTGGCTGTTCCGTTGGGCTACTT 2910  
 Qy 2757 aattcaagctgagaaagagacccagagagcatgtctatagtttactgcatgtgata 2816  
 Db 2911 GGTGCAATATAGTTATGATGAACTTTCAAGAGATGTTCTTTGGATTTCTAGCATATGCTT 2970  
 Qy 2817 tgaagctatgagcagtaaaagcgtcttgcctcctgaatataccaagccaatc 2876

Db 2971 TGTGTCATGTCAGTGGGAGCAGTCACTTATTTGCTCTGACTATGCAAAAGCCAAAGT 3030  
 Qy 2877 gggggtgcgcatctgttgccttctgttggaaagaagaacaaatataagacgcgcatga 2936  
 Db 3031 ATCAGCAGCCACAGTCATCATGATCATTTGAAGAAAGCCCTGATTTGACACCTACAGCC 3090  
 Qy 2937 agaagggaaagccagacacatgtgaaggaattagattcgtgaagctccttctt 2996  
 Db 3091 TCACGGCTCAAGCCAAATACCTTGAAGGAATGTGACATTTAATGAGTGTCTTCA 3150  
 Qy 2997 ctatccatgctgcagagatgttccatcctcgttgcgttccatccctcaatgattgagcag 3056  
 Db 3151 CTATCCACCTGACACAGACATCCCGTCTCCAGGCGCTGAGCCGCGAGTGAAGAAAGG 3210  
 Qy 3057 aaagaagtagcatltgttgggagcagcgtctgtggaaagacatctgttcaacttc 3116  
 Db 3211 CCAGACGCTGCGCCCTCGAGTACAGTACGCTGTGGGAAAGACACAGTTTCTTCACTCT 3270  
 Qy 3117 gcagagacttataagcccgctgcaagagagcatgtctgttgaatgtgtgagtgcaaaaga 3176  
 Db 3271 AGAGGCTTCTATGACCCCTTGGCTGTTCAGTCAATTAATGAGCAAAAGATTAAGCA 3330  
 Qy 3177 attgaatgacagtggtcgtctcccaatagcaactgttcccaagagcctgtcct 3236  
 Db 3331 CCTGATGTCCAGTGGCTCCGACGACACCTGGGATGCTGTCTAGAGGCCATCTGTT 3390  
 Qy 3237 caactgcaagatgtctgagaacatcgctatgtgtgaacaagcgtgtgtgtccattaga 3296  
 Db 3391 TGACTGCGCATTTGCGGAAACATTTGCTATGAGAACACAGCGGCTGTATACATGA 3450  
 Qy 3297 tgagatcaagaagccgcaaatgtgcaaatatccatctttaaagagtcctcctga 3356  
 Db 3451 AGAGTTTTCAGAGGACGCAAGAGGCCACATACATCACCTTTCATGAGACATCTCCGTA 3510  
 Qy 3357 gaatatacaacacacagtttgacttgaagaagcagcgttctgtgcgcagaaacaaag 3416  
 Db 3511 GAAATACACACACAGATGAGAGAACAAAGGAACCCAGCTCTGTGGTGGCCAGAAACAGCG 3570  
 Qy 3417 actagctatgcaagaggtctcttccaaaacccaatttatgttgtgtgagggcac 3476  
 Db 3571 CATTCGCATACCTGCGCTCTTGTAGACACCTCATTTTGTGCTTTGGATGAAGCTAC 3630  
 Qy 3477 ttgagccctgataatgacagtgagaagtggttcaagatcccttgaataagagagac 3536  
 Db 3631 ATCAGCTGTGATACAGAAAGTGAAGAGTGTCCAAAGACCTGTGACAAAGCCAGAGA 3690  
 Qy 3537 ggaagagacatgctagctgacatccacagagcctcctgcaattcagaacgcagattgat 3596  
 Db 3691 AGGCGGACACGTCATGTGATGTCGCCACCGCTGTGCCACATGCCAGATGCAATTTAAT 3750  
 Qy 3597 agtgttctgcaaatggaagataaaggaacgaacatcaagaagcctcctgagaa 3656  
 Db 3751 AGTGCTGTTTCAAGATGCAAGATCAAGAGCATGAGCATGCAATCAACAGCTGTGGCCA 3810  
 Qy 3657 tcagacatatattttaaagttagttgaatgacagcag 3694  
 Db 3811 GAAAGCATTAATTTTCCATGCTCAGTGTCCAGGCTG 3848

RESULT 2  
 AX105080 4279 bp DNA PAT 30-APR-2001  
 LOCUS AX105080  
 DEFINITION Sequence 24 from Patent W00123540.  
 ACCESSION AX105080  
 VERSION AX105080.1 GI:13921230  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE  
 1 (bases 1 to 4279)  
 Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and  
 Patten,C.J.



OY	1719	tgyaaycctgagcgagaaagagacatgctgtaactaatgycanaacgagtgctatalt	1778
Db	1831	TGGAGTCATTGTTGGAGAAAGAAATATCATGTATGAACTCATGAAGAAGAAAGGCGCATTTACTT	1890
OY	1779	tcgaactgtgatgycacaagatbtctaaaaagctgtga--acaagtgyagtcataagcat	1837
Db	1891	CAAACTGTCTCAATGACGACCAAGAGAAATGAAATTTGATTAGAAATGCGCATGCTGTGA	1950
OY	1838	attctactgaaagaagacccaactcactccctctgcaactctgtgaagag-----cat	1889
Db	1951	ATCCAAAGTCGAAGTATGATGTCCTTGGAAATGTGTCCAAAGATTCAAGGCTCCAGTTTAT	2010
OY	1890	caagtcagacttaattgacaaagcctgagaatccaccaatctaaagatgaagtcct--	1947
Db	2011	AAAAAGAAATCAACTCGCAGAGATTTACATGCACCACCAAGGCCAAGACAGAAACTTGG	2070
OY	1948	-----cctgaagctctctcattataaaatttaagtt	1979
Db	2071	TACAAAAGAGACTTGAATGAAATGTACCTCCAGTTTCTTCTGTGAGAGATTCTGAACCT	2130
OY	1980	aaacaagcctgaatgagcctttgtgtctcgtggagcaattgsgcttcgttctcaatgagac	2039
Db	2131	GAACTCAACTCAATGAAAGCGCTTATTTTGTGTGTGTATTTTGTCTATTATTAACGGAGG	2190
OY	2040	tgattccagtaatttccatcactcttgcnaaaatataaacaatgltt--ggaataa	2096
Db	2191	CCTGCACACGACATTTTCAATATATTTTCAAGATATTAGGAGTCTTTACCCGAGATGA	2250
OY	2097	tgataaacaacatlaaagcatgagtcagaaattatccatgatatactgcataattggt	2156
Db	2251	GGATCTCGAAACAAACAGACAGATAGTAAACATGTTTCTGTATTGTTTCTAGTCTTGG	2310
OY	2157	tgtaatttgcctgtgcagatattctcagaggaattatttcagcagaagagggaat	2216
Db	2311	AATTATTTCTTTTATATCAATTTTCTCCAGGGCTTCACATTTGGCAAGCGTGGGAGAT	2370
OY	2217	tttaacgatagatlaaagacacttgsccttcaaaagcactgtatatacagatattgscgt	2276
Db	2371	CCTCATTAAGGGCTTGATACATATGGTTTTCACATCCATGCTGACGACAGATGTCAGCTG	2430
OY	2277	gtttgatgtaaaagaaacagcagcagagagcttgcacaacaatatatagcatagatatagc	2336
Db	2431	GTTTGATGACCTTAAACAAACCACTGGAGCATTTGACAAACAGGCTTGCCTCAATGATCGCG	2490
OY	2337	acaaattccaagagcgaacagcttccagatitgscgtcttaacacaaatagcaactaacat	2386
Db	2491	TCAAGTTAAAGGGCTTATGAGTTCACAGGCTGTGTCATTTACCCGAAATATGCAAACT	2550
OY	2397	ggagacttcaagttcatcttccctctatatagtatgtggagatagcaatccctgattctgag	2456
Db	2551	TGGGACAGGCGATTATTTATCTCTTATCTATGTTTGGCAATTAAACATTTTACTTTAGC	2610
OY	2457	tattgtccagttactgctgacgaagatgaltgaacccgcaagcaatgcatgcatgttc	2516
Db	2611	AATTGTACCATCTATTCATATACAGAGAGTGTGTAAGAAATGTTGTCTGACAGAC	2670
OY	2517	caacaaagataagaagaacttaagcatgctgtgaaagataagcaactgaagcettgagaa	2576
Db	2671	ACTGAAGAATTAAGAAAGGCTTAGAAGAGCGTGGGAATTTGCTAACGAAGCCATCGAATA	2730
OY	2577	tatacgtactaagtggtcacttaacgaaggaanaagccttcggagcaatgtatgaagaagat	2636
Db	2731	CTTCCGAAGCTGTGTTCTTTGACTCGGGACGCAAGACTTGGATATACATATATGCAACAG	2790
OY	2637	gcttcagactcaacacagaataatacctcogaagaaagacacagatatttgaagctgtatgc	2696
Db	2791	TTTTCAGATACCATACAGAAACTCTTTGAAGAAAGCACACATCTTCGGGTCTCATTTTC	2850
OY	2697	atcaagcgaatgcttatataatttgcctatgtcagaaggttgcattttgaagcctattt	2756
Db	2851	TATACACGAGCATGATGATTTTTCATATGTCTGGCTTTTCCGTTGGTTGGCTCTACTT	2910

OY	2757	aatcaagctcggacgaatcgcgcccaagagcgcaatgltcaagtlttactctgcaattgcta	2816
Db	2911	GGTGGCAAAATGAGTTTCATGTAACATTTTCAGGATGTTCTTTTGGATATCTCAGCATTTAGTCTT	2970
OY	2817	tgagagcatcgccacatcgagaaacagctcgtcttggtcctcgatattccaaagccaaac	2876
Db	2971	TGGTGCCATGGCACTGGGGGCAAGGTCAAGTTCAATTTGCTCTCTGACTATGCCAAMGCCAAMGT	3030
OY	2877	gggggctcgacatcgtctgttcctgtctgtgaaaaaagaaaccaaatatagacgcgcagtica	2936
Db	3031	ATTCAGCAGCCCACTGATCATATGATTCATATGAAAAAAGCCCTCGATATGACAGCTACAGCCC	3090
OY	2937	agaaaggaaaaagccagacacatgctaaagggaaattagaagtttcgagaagctctcttct	2996
Db	3091	TCACGGCCCTCAAGCCCAATATCGTTGGAGAAATGTGACATTAATGAGAGCTGCTTTCGA	3150
OY	2997	ctatcacatgtcgcccaatagtlttcactccgtcggtcctatccctcagatcagtcagcgag	3056
Db	3151	CTATTCCACTGACCAACAATCCCCGCTGCTCCAGGGGCTGAGCCTCGAGGTAAGAAGGG	3210
OY	3057	aaagacagtagcaattctgtgggagacagcgctgtgtgaaaaagcaactctgtcaactctc	3118
Db	3211	CCAGACGGTGGCCCTCGTAGTGAAGTGAAGTGGCTGTGGGAAGACACAGTGTTCACCTCCT	3270
OY	3117	gcagagacattatgaacccgctgcagaagacaagtgctgttttagtgbtgbatgataaaga	3176
Db	3271	AGACGCGCTTCTATATACCCCTTGAGCTGAGTTCAGTCAATATGAGCAAGAGATATACCA	3330
OY	3177	attgaatgtacagtggtctcgttccccaataagaaatcgtctctcaagagccgtgtctct	3236
Db	3331	CCTGAATGTCAGTGGCTCCGACACACCTGGGCATCTGTGCTCAGAGACCCATCTCTTT	3390
OY	3237	caactgcagcatctcttagaacaatcgactatgtgcacaacgctgtgtgtgcataaga	3296
Db	3391	TGACTCGACGATTTCCCAAGAACTATGCTTATGAGACACACCGGGTGTATCACAATGA	3450
OY	3297	tgagatcaagaagccgcgaatgcagcaaatcaltcatctttatgtaaagtgctccctga	3356
Db	3451	AGAGATTATTCAGGCAAGCCCAAGGAGGCCAATACACCACTTCATGAGACACTCCCTGA	3510
OY	3357	gaatatcaacacaaagtctggacttgaaagagacacgcttctctgscgcgcgaataacaa	3416
Db	3511	GAATATACAAACCCAGAGTATAGAGACAAAGGAAACCCAGCTCTGTGGTGCCAGAAACACG	3570
OY	3417	actagctattgcgaagcgtctctctccaaaaccaaattttattgtgtgatgagcgac	3476
Db	3571	CATTGCCATAGCTTCCGCTCTTGTGTAGACAGCTCATATTTTGTCTTTTGATGAAGCTAC	3630
OY	3477	ctcagccctcgatatactacagtgagaaagtggtgttcagcatgcgcctgtataaagccagac	3536
Db	3631	ATTCAGCTCTGATATCAGAAAGTGAAGAAAGTTGTCCAAAGAGCCCTGAGCAAGCCAGAGA	3690
OY	3537	gggagaagacatgctagtggtgactcaatcaagagctctctcgaatcagaacgcagattgat	3596
Db	3691	AGGCGCGCACCTGCTGTGATGTCGCCACCCCTTGTCCACATTCAGAAAGCAGATTATAT	3750
OY	3597	agtggtcttcgacaatggaaagatlaaagggaagaagaaactcgaacagctctcgtcgaaa	3656
Db	3751	AGTGGTGTTTAGATATGGCAAGTCAAGGAGCATGGCACAATCAATCAACAGCTCTTGCCCA	3810
OY	3657	tcgagacatatatttaagttagtgaatgaaatgcagatcag	3694
Db	3811	GAAAGGCACTCTATTTTTCATGTGTCAAGTCCAGGCTG	3848
RESULT	3		
AX105082	AX105082	4279 bp	DNA
LOCUS	AX105082	Sequence 26 from Patent WO0123540.	PAT
DEFINITION	AX105082		30-Apr-2001
ACCESSION	AX105082.1	GI:13921232	
VERSION			
KEYWORDS			
SOURCE	dog.		

ORGANISM      *Canis familiaris*  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE      1 (bases 1 to 4279)  
 AUTHORS      Stocker, P.J., Steelmeier-Crespi, D.T., Crespi, C.L., Reif, T.C. and  
                   Patten, C.J.  
 TITLE            P-glycoproteins and uses thereof  
 JOURNAL        Patent: WO 0123540-A 26 05-AR-2001;  
 FEATURES      GENTEST CORPORATION (US)  
                   Location/Qualifiers

```
/organism="Canis familiaris"  
/db_xref="taxon:9615"  
17. .3862
```

BASE COUNT 1296 a 833 c 1009 g 1141 t  
ORIGIN  
/protein\_id="C9C37723.1"  
/db\_xref="GI:13921233"  
/translation="MDPEGGKRGKSGAEKNFMWKKSKKKKKKKPTVSTFAMFRYSN  
MLRRLYLVEGTMAIINGAALPIMLALVFGMTQSFANAAGISRRKTEPVINIEETINTQ  
OHLFNHLEEMTAYAYYSGISGALVLAATVQSFVLAAGROI LKRFQFPAHIMM  
EIGEFVDHVGELNLTRLTDDYSKINESIGKIGMFOSITPFPGFVITRGKKTLL  
VILIASPVILGSALIAIMAKITITSEPKDELLVAAKGAEEVILAIPIVIAFGOKKEL  
EPRKNLNEAKGIGICIKKAITANISIGAFLLITASYALAFWGTSLVLSSEYITIGOYL  
TFEVSVLIGAFSIGQASPSIEAFANAKGAAYELFKIIDNPSIDYSKSHKPDNKL  
NLEKRNHGFYSFSPKRKEXKILKGLNKLKQOSQATLVNCGSGSTVOLQKRPIDPD  
GAVICDIOGDIRITNVRHLREITGVSOPELPAITLAINTRGENYTMDEIKAEVW  
AMADIPIMKLPNKEDPTLVGBRGALDSGOKORAIALVRNKGITILLDPAATLDE  
SEANVOYALDKARKGRRTIYIAHRLSTVRNADVIAGDDCVIYENKHHMLMEKGIY  
FKLTWOTRGNEIELEENATESSKSESALMSPDSSSLIKRKPSRTHIAPOGQOR  
KLGEEDLNEBVPVPSFWRLIKLUNSTPEWPFVVGIFCAIINGGLQAFSTISFRIJGI  
TRFDEBEPKRONSNMPSVFLVGLIISTFTFLOGTFEKGAGILTRKRLRYAFISM  
LPROBSEWDDPKNTGTGALTRLANDAQAQVGAIGSRALVATTONANIGTGLIISLJG  
MQLLILLAIYPIIAGVEMKMLSGOALDKKKELEGACKITAEONLEBRTVYSLTR  
EOKREYVTAOSLOYPIYRNSLIRKAHIFVSSITQAMNTEYIAGCFRGATLVANEPNR  
FOVDLVFSAIVFGAMAVGVSSFPAPYAKARSAAHVIMIEKSPILDSYSPHGKLP  
NLEGTANTFNEVEYFNPTRPDIPLYGLQISLEVKQGLTALVSGSGCKSTVYOLLEEF  
YDPLAGSVLIDGKEIKLNTQWMLRAHIGIOWEILDCSIAENIAGDNRVSVHEE  
IYQAKENAHNHIEITLPEKRYNTPVSGKQGLSGOKORAIARAIVRQPHIILLDA  
TSALDESEKVOYDALDKAREGRCTIYIAHRLSTIQNADLIYVQNKVKEHGHQOL  
LAQGIITFSMVSQAGAKR"

Query Match	33.3%	Score 1234.4	DB: 6	Length 4279
Best Local Similarity	61.2%	Pred. No. 1.5e-247		
Matches 2190	Conservative	0	Mismatches 1256	Indels 132
			Gaps	
OY	162	caggttcagccctgataatctgttcgaataagatgttcctgccttgatttgygtacataca	221	
Db	358	CACGATATGCGCTATTTATTTACATGGGATGCGTCTGGCGTGGTGGCGCTTACATCA	417	
OY	222	gatttccttgygtattataaactgcagcacagacagaccagagattcgaanaacagtttt	281	
Db	418	GCTTTCATTCTGTGGTCCCTGGCAGCAGGAAGACAGATACTCAAAAATTAGAAAATTTT	477	
OY	282	tcattcagttttggcacagggacatctgctgttttatagctctgtgacatcgtggaactaa	341	
Db	478	TCATCTATTCATGCGCACAGGAGATGTGGCTGGTTTACGTGATCGCTGGGGACCTTAA	537	
OY	342	cactgcagatgacagacatgy--acaaaatcagtatggtatgtgagaataattgcctc	398	
Db	538	CACCGCGCTCACAGACGATGTCTCCAAAATCATATGAGGAATTTGGCGCAAAAATTGGAA	597	
OY	399	gttgtttcaaaaacatgctcactcttcttcgatattgacctgacgttggtttggtaaggtct	458	
Db	598	GTTCTTTTCATCAATAGCAACATTTTTCACCGGTTTTATATGGGGTTTACACGTGGTTG	657	
OY	459	gaaactaacccctcagtgacctctatccaagctctcctcttataatggtcttaagcgacgcatg	518	
Db	658	GAAAGTAAACCTTGTGATTTTGGCATCAGACCTGTGTTTGACATTTGACCGGCATCTGG	717	

OY	519	ttctagaatgtaatcctctctgacccgaatgaagaaatlaagtgcttaatccatcccaagcttgggc	578
Db	718	GGCAAAATATCTATCTTCATTTACTGATTAAGAAGCTCTTGGCCTTATGCAAAAGCTGGAGC	777
OY	579	tgtgacgaagaagctctgtcatalcaatccgaacagtcataagccttlaagggccagagaa	638
Db	778	AGTAGCTGGAAGAGTCTTAGACACATCAGAACTGTGATTTGGCTTTGGAGGACAAAGAA	837
OY	639	agaactccaaggcttccctttaataataacaagataigtggttttaatttcccaca	698
Db	838	AGAAC-----	842
OY	699	gtggctactaagttggtctgttcttttgtaagratcaacgaatccaaatgcgaatgcaaa	758
Db	843	-----TTAAAGGATACACAAAATTTAGAGAGACTAA	876
OY	759	ggaattgacataaaagagactaagcttccaaagtgctctgtgtcgtgtaactctt	818
Db	877	AGGAATTTGGGATTAAGAAGACTTACAGGCCAAACATTTCTATGTGTGGCGCTTCTTATT	936
OY	819	tatgaatgaacctaagagactgtcttttgtaatgaagaacctctgaattcttaatgaga	878
Db	937	GATCTAGCATCATATGCTCTGTGGCTTTCGTGATGAGGACCTCTTGCGCTCTCAG--	993
OY	879	accggatatacactcggagactgtctgtctgtctttcttttagtgaatccatalagcagta	938
Db	994	--TGAATATACATTTGGACAGATCTCACTGCTCTTCTTGTGATTAATTTGGGCTTT	1050
OY	939	tgtcattggaagcagcagctccctcactcttgaacacctgcgaatagccgagagagctcct	998
Db	1051	TAGATTTGGACAGGACATCCCAAGCATTTGACGATTTGCAAAAGGACAGAGCGTTA	1110
OY	999	tcatatttccaggttattgataagaaccagatalagaaactttccacagctgata	1058
Db	1111	TGAATCTTCAAGTATTAATTGACATTAACCAAGACATTCAGACATTTGGAAGAGTGGACA	1170
OY	1059	taaacctgaatcatagaaagaaactgtggaatttaaaatgcttcttcaattcatc	1118
Db	1171	TAAACCGATATATTTAAGGGAATTTGGAATTTCAAAATGTTCACTTCACTTACCTTC	1230
OY	1119	aagcacatcatcaagaattctgaaagcttcgaatccagaattlaagtccttgagagacagt	1178
Db	1231	TCGAAAAGAAATTAAGTCTTAAGAGGTCTCAACCTGAAGGTTGAGATGGGGACAGCT	1290
OY	1179	cgccttgctcgttccaatgagcagtgaggaagtagtaagtagtccagcttctgcagagtt	1238
Db	1291	GGCCCTGTTGGGCAACGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGAGTCGAGAGCT	1350
OY	1239	ataatgaatccggaatgagcttatcatgtgtgtagaagaaatgacatcaagatttaaatgt	1298
Db	1351	CTATATGCCCCACATATGGCATGTGCTGTATTTGATGAGACGGGACATTAAGACCATTAATGT	1410
OY	1299	ggcgcatattcagaaccataattgagtggttagtcaagaagcctgttcttgcctgcggaacac	1358
Db	1411	AAGGCATCTTGGGAATTAATTACTGTGTGGTGGATGAGACGCTGTGTTGGCACACAC	1470
OY	1359	catcagtaacaatatcaagatagagacagagatgtyacatgataagaagatgtagagagc	1418
Db	1471	GATAGCTGAAAACATTCGTATGCGCGCAAAATGTCACCATGATGATTTGAAGAAAGC	1530
OY	1419	agcaaggaagcaaatgcytatgatttatcatatgagaqtttctaataaatttaactt	1478
Db	1531	TGTTTAAGAGGCCAATGCCATATGATTTTATCATGAAAACATCAATTAATTAATTTGACACTT	1590
OY	1479	gtgtagggaaaaaagagctcaaaatgagtgtagagagcgaaagaagagatccgaattgtcg	1538
Db	1591	GGTTGGAGAGAGAGGGCCCAAGCTGATGAGTGAGACGAACAGAAATTCGCCATTGGCTCG	1650
OY	1539	tgccttgatttgaaaaccccaagattctgattttagatgagctlaagctctgcctcggatttc	1598
Db	1651	GGCCCTGTTGGCAACCCCAAGATTTCTTCTGCTGATAGAGCAAGCTCAGGCTGGAGAC	1710
OY	1599	agaagaagatcagctgttccaagctcgaactggaagagcgagcaaaagttcggactacaat	1658

```

Db 1711 TGAAGTGAAGCAGTGGTTCAGTGGCCCTGATTAAGGCAGAAAGGCCGACTACAT 1770
QY 1659 cgtggttagcaaccagctcttctactatctgaagtcgaagttgattgtgacccctaaag 1718
Db 1771 TGGATGATGCTCATCGTTTGTCTACAGTTCGTAAATGCCGATGTCGTGTTTATGA 1830
QY 1719 tggaaatgctgaggaagaagcacatgctgaactaagtcgaagaacgagttctatca 1778
Db 1831 TGGAGTGCATTTGTGAGAAAGAAATCATGTGAACTCATGAAAGAGAGGCAATTACTT 1890
QY 1779 ttcacttgcattgacagagatactaaagctgtga-acagatggaatcgaatgacat 1837
Db 1891 CAACCTTGTCAATGACAGCAAGAGAAATGAATTTAGATTGAAATCCACTGCTGA 1950
QY 1838 atctctctgaagaagaaccaactcctcctcctgaactctgtgaag------cat 1889
Db 1951 ATCCAAAGAGGAAAGTGAATGCCCTTGGAAATGTCTCCAAAGATTCAGGGTCCAGTTTAA 2010
QY 1890 caagtcgaacttcattgacagagctgaggaatccaccaatctaaagagtaagttct-- 1947
Db 2011 AAAAAGATCAACTGCGAGGAGTATACATGCAACAAGGCCAAGACAGAAAGCTTGG 2070
QY 1948 -----cctgaagtcctcctaataaaatttaagtt 1979
Db 2071 TACAAAGAGGACTGTGATGAGAAATGACCTCCAGTTTCTCTTGAGAGATTCGAAGCT 2130
QY 1980 aaacaagcctgaatgagccttctgtgltctgggaactgagctctgtcttaaaatggaac 2039
Db 2131 GAACCTAACGTAATGAGCTTATTTTGTGTTGTTGATATTTTGTCTATTTTAAACGAGAG 2190
QY 2040 tgtctacgaattcttccatctcttgcataatataaccatctt---ggaataa 2096
Db 2191 CCGCAACCCAGCACTTTTCAATATATTTTCAAGGATTAAGGATCTTTACCCGAGATGA 2250
QY 2097 tgaataaacacataaagaatgacagaatcttcaatgatatctgtcattgtg 2156
Db 2251 GATTCCTGAACAAAGAGAGAAATGATACATGTTTCTGTATTTGTTTCTAGTCCCTGG 2310
QY 2157 tgtatttgccttgcagttatcttcatctgacaggaatcttcaagcagagagaggaat 2216
Db 2311 AATTATTTCTTTTATTAACATTTTTCCTCCAGGGCTTCACATTTTGCATAATTCAGAACCT 2370
QY 2217 ttttaagaatgataaagacacttgccctcaagaacatgttatacagaatattgctg 2276
Db 2371 CCTCAGTAAGCGCTTCGATACATGTTTTCAGATTCATGCTGAGACAGAGATGTCAGCTG 2430
QY 2277 gtttgaatgaagaagaacacagcagagagctgacacacatactagccaatagatagc 2336
Db 2431 GTTTGATGACCTTAATAAACACACTGAGCATTTGACACACAGGCTTGCCAAATGATGGCG 2490
QY 2337 acaaatcaagaagcaacagttccagagattgcgctcttaacaacaatgcaactaacat 2396
Db 2491 TCAAGTTAAAGGGGCTATAGTTTCCAGGCTTGCTGTCATTATCCAGAAATATACCAATCT 2550
QY 2397 gggacttcaagtaatacttcccttatalatgtagtggagatgaactccttgctgag 2456
Db 2551 TGGGACAGGCAATTAATATCTTAATCTATGTTGGCAATTAACACTTTTACTCTTAGC 2610
QY 2457 tatgttccagtaactgacgctgacaggaatgaaacccagcaatgactgatttgc 2516
Db 2611 AATTGTACCCATCATTTGCAATAGCAGAGATTTGTAATTAATGTTTGTCTGGACAGC 2670
QY 2517 caacaaagaatgaagaactaagcattgactgtgaagaatagcaactgaagcttggagaa 2576
Db 2671 ACTGAAGATTAAGAAAGAGTGAAGAGACTGGGAAGTTGTACAGAACCATTCGAAAA 2730
QY 2577 tatacgtactatagttcattcaacaagaagaagccttcaggaacaaatgfatlaagagat 2636
Db 2731 CTTCGGAACGTGTTGTTCTTTGACTCGGAGCGAAGATTGGAATACATGATGACAGAG 2790
QY 2637 gcttcagactcaacagaataactclogaaagaagacagatattggaagctgtatgc 2696

```

---

```

Db 2791 TTTGCAAGTACCATACAGAAACTCTTTGAGGAAGACACACATCTCGGGGTCTCATTTTC 2850
QY 2697 attcaacatgctcttatalatttgcctatgagcaaggttccagtttggagccattt 2756
Db 2851 TATCACCCAGGCAATGATATATTTTCTATGCTGCTGTTTCCGGTTTGGTCTACTT 2910
QY 2757 aattcaagctgagcaatgaccccaagggcalttcatatglttctactgcaattcata 2816
Db 2911 GGTGGCAATGATGATGATGAACTTTTCAAGATGTTCTTTTGTATTTCTCACTATTTGCT 2970
QY 2817 tggagctatggccalcggaanaacgctcgttctgctcctgaataltccaagcaaatc 2876
Db 2971 TGTGTCATGCGAGTGGGGCAGGTGATTCATTGCTTCCCTGATGATGCCAAAGC 3030
QY 2877 gggggtctgcacatctgttgccttctgtggaanaagaacaaatatagcaagccgagta 2936
Db 3031 ATCAGAGCCCAAGTATCATGTATGTAATAAACCCCTCTGATTGACAGCTACAGCC 3090
QY 2937 agagggaaagacacagacacatgtagaaggaattagagtttcgaagctctcttct 2996
Db 3091 TCAAGGCTTCACCAATATGCTTGAAGAAATGTGACTTTTAATGAGTCCGTGTTCAA 3150
QY 2997 ctatccatgctgcacagatgttctacatcctcgttgccttaccatgattgagcgaag 3056
Db 3151 CTATCCCACTCGACAGACATCCCGTCTCAGAGGGCTGAGCTCGAGGTGAAGAGG 3210
QY 3057 aaagacagtagcaattgttgggagcagcgctgttgggaagaagccttctgttaacttct 3116
Db 3211 CCAGAGCTGCGCCCTGTAGTGTAGAGTGTGGAAGAGCAAGTTGTTAGTCTCT 3270
QY 3117 gcagagacttataagacccgtgcaagaacagctgttgcattgagctgtgtagcagaag 3176
Db 3271 AGAGCCCTTATGACCCCTTGCTGCTGTTCAAGTCTAATTTGATGCGCAAGAAAGTAAAGCA 3330
QY 3177 attgaatgacagtggtcctcgttcccaaatgaatcgttccctcaagagcgtgtcctt 3236
Db 3331 CTGAATGTCTCAGTGTGCTCGAGCACACCTGGGCACTGTTGTCAGAGAGCCATCTCTGTT 3290
QY 3237 caactgacagatgtctgagagacatgcctatggtgcaagaagcggtgtgctcattga 3296
Db 3391 TGACTGAGCATTTGCCAGAACATTCCTATGAGAACAAAGCGGCTCTATACATGA 3450
QY 3297 tgaatcaagaagccgcaaatgcaagcaaatccatcttcttatttgaagctccctga 3356
Db 3451 AGAGATTTGTCCAGGCAAGCAAGAGGCCAATACATACATTTATGAGACACTCCCTGA 3510
QY 3357 gaaatacaacacacagttgagctgaagaagcagcttctgcgagccagaacaag 3416
Db 3511 GAATATCAACACAGATAGTAGAACAAAGAACCCAGCTCTGTGAGGCCAGAAAGCG 3570
QY 3417 actagctattgcaaggtcttctccaaaaaaccaaatlttatgttgaatgagccac 3476
Db 3571 CATTTGCATATGCTCGGCTCTTTGTAGACAGCTCATATTTTCTTTTGGATTAACCTTAC 3630
QY 3477 ttcaagccctcgataatgagagatgagaggtgttgcagctgagccttgataaagcagagc 3536
Db 3631 ATCAGCTCTGATACAGAAAGTGAAGAAAGTTGTCCAAAGAACCCCTGACAAACAGAGAG 3590
QY 3537 gggaaagacatgctagttgttcaatcacaaggtctctgcaatctcgaagcagatltgat 3596
Db 3691 AGGCGGACCTGATGATGTATGTCGCCACCGCTTGTCCACATCCAGAAATGCAATTTAAT 3750
QY 3597 agtgtgtctgcaaatggaagataaagaagaagaacatcattcaagaagctccttgagaa 3656
Db 3751 AGTGTGTTTCAAGAAATGCAAGAGTCAAGGAGCATGTGCACATCAACAGCTGCTGCCCA 3810
QY 3657 tcaagacataatttaagttaagttgaatgacagctcag 3694
Db 3811 GAAAGCATATTTTTCATGCTGCTCACTGTCAGGCTG 3848

```

RESULT 4  
AX105057



[illegible]

Db	558	GTCTTTACTCAATACCAACATTTTTACCCGGTTTATAGTGGGGTTTACACGTGGTGG	657
Qy	459	gaactaccccttagtgactcatcacacgctccctcttataatgcttcaagcggcagcatg	518
Db	658	GAAGCTAACCTTGTTGATTTTGGGCATACAGCCCTGTGTTTGGAATTTACGGCCGACATCG	717
Qy	519	tcttagatggtcatcctcaatcagcaagaagaatgaatgctctatccaaagctgggc	578
Db	718	GGCAAGATTAATCTTATCTTCAATTTACTGATTAAGAAACCTTTGGCCATGAAGAAAGCTGGAGC	777
Qy	579	tgtggcaagaagaatctctgcatcatccgacaagcagcatagcctttagggccagagaaa	638
Db	778	AGTAACTGAAAGAACTCTTAGCAGCAATCAGAACTGTGATTTGCTTTGGAGGACAAAGAA	837
Qy	639	agaacttcaaaagtccttcctctttaaataaacaagatagcttggttllatltcccca	698
Db	838	AGAAC-----	842
Qy	699	gtggctactaagtgtgtgtctgtnttltglaaglatacacagaaatcctcaaaatgycaa	758
Db	843	-----TTGAAAGGTACACAAATAATTTAGAAAGAGCTAA	876
Qy	759	ggatttggcctaaagaagcctatagcttcaaaagtgctctctgtgtgtgtactctt	818
Db	877	AGGAATTTGGATTAAGAAAGCTTATCACGGCCACATTTTATATGGTGGCCGCTTCTTATTT	936
Qy	819	tatgaatgaacctatgacttgcttcttgttgaatggaaacctctgtatcttcaatgagaa	878
Db	937	GATCATATGCATCATATATCTGCTGGCTTCTGTGATGGAGACCTCTTGTCCTCTCCAG---	993
Qy	879	accttgaatataccatcggagacgtgtctgtcttcttctttagtgaatcatatgacgtta	938
Db	994	---TGAATATTTCTATTGACAAAGTACACTCATGCTCTTCTTGTATTAATTTGGGGCTTT	1050
Qy	939	tgtcatttggagcagcagctccctcaatcttgaaccttcgcaatagccgagagagctgctt	998
Db	1051	TAGTATTTGGACAGGCATTCGCCAAGCATATGAGCATTTTGCAAAAGCAAGAGCAGCTTA	1110
Qy	999	tcatatttccaggttatatagaagaaccagatagataaacttccacagcctgagata	1058
Db	1111	TGAATCTTCAAGTATTTTGCATATAACCAACGATTTGACACTTATTTGGAAGATGGACA	1170
Qy	1059	taaacctgaatccatagaagaacgtgynaatttaaataatgcttcttcaatlatcaatc	1118
Db	1171	TAAACAGATATATTTTAAAGGAATTTGGAAATTCAAAATGTTCACTTACGATTCACCTTC	1230
Qy	1119	aagacatcatcaagcttctgaagaagcttgaatctcgaatataagcttgaagagacgt	1178
Db	1231	TCGAAAGAACGTTTAAAGTCTTTAAAGGTCTTCAACCTGGAAGGTTCCAGATGGGCAACACGT	1290
Qy	1179	ggccttgtgtgcttcaatgacgttgaagaagatacggtagtcagcagcttctgcagagatt	1238
Db	1291	GGCGTGTGTTGGAAACGTGGCTTCGGGAAGACACGACCGTGCACGTGATGACAGAGCT	1350
Qy	1239	atatatcccgatgatgtcttcatatgtytgaatgaaatgacatcagcagagctttaaagt	1298
Db	1411	AAGGATCTTTCGGGAAATTTCTGCTGTGGTGAATGAGAGCTTGCTGTTTGGCCACACAC	1470
Qy	1359	catcagtaacaatatcaaatgataagcagagatgtagtgaatgataagaagatggagaagc	1418
Db	1471	GATAGCTGAANAATTTGGCTATGGCCGCGAANAATGTCACCATGATGAGATTGAGAAAGC	1530
Qy	1419	agcaaggaagcaaatgcatgatatttcatatcagtgaagtttccctaataatttaacat	1478
Db	1531	TGTTTAAGGAACCAATCCATATGATTTTATCATGAANAATCTAATTAATTTGACACTCT	1590
Qy	1479	gttaggggaaaagagcttcaaatgagtggaggcagaaacagagatccgcaattgtctg	1538
Db	1591	GGTTTGAGAGAGAGGGGCCAGCTGATGTGTGTGACGAANAACGAGAAATTCGCATCTGCG	1650



OY	1539	tgcccttagtcgaaccccccaagttctgattttagatgtagctgaagtcgtcccttgatgc	1538
OY	1539	tgcccttagtcgaaccccccaagttctgattttagatgtagctgaagtcgtcccttgatgc	1538
Db	1651	GGCCCTGCTGGCAACCCCAAGATTCTTCCTGCTGGATGAGGCAACGTCAAGCTTGGAAC	1710
OY	1599	agaagaagaatgcacgtctttaaagctgcacgtgcgaagaagcgagcaaaagtcgaactaacat	1658
Db	1711	TGAANAATGCAAGCAATGGTTTAAAGTGGCCCTGGATTAAAGGCCAATAAAGCCGAGACTACCAT	1770
OY	1659	cggtgtagcaacacccgactttctactatctcgaaagtcgaattgattgtagccctaaaga	1718
Db	1771	TGATATGCTCATCGTTTGTCTACAGTGTGTAATGCGCATGTCAATTTGCTGGTTTGATGA	1830
OY	1719	tggaatgctgcgcggaagaagagcaaatgctgaactaaatgagcaaaacgaatgctatat	1778
Db	1831	TGGAGCTATTGTGGAGAAAGAAATATGATATGAAATCAATGAAAGAAAGAGCGCATTTACTT	1890
OY	1779	ttcaactgtagatgcacaagatalataaaagctgtatga-acagatggagtcgaatgacat	1837
Db	1891	CAAACTTGTCACAAATGCAAGCAAGAGAAATGAAATGAGTTAGAAATGGCACATGGTGA	1930
OY	1838	attctactgaaagaagaaccaactcacctctcctctgcactctgtgaagag-----cat	1889
Db	1951	ATCCAAAAGTGAAGTGCATCCCTTGGAAATGTCTCCAAAAGATTCAGGCTCCAGTTTAAAT	2010
OY	1890	caagtcgaactcaatgcacaaggtctgaggaatcaccacaactaaagaatgaagctt--	1947
Db	2011	AAAAAGAGATCAACTCGCAGGAGTATACATGCACCACAGGCCACAGACAGAAAGCTTGG	2070
OY	1948	-----cctgaagctctcctatataaaattttaaagt	1979
Db	2071	TACAAAAGAGACTTGAAATGAGATGATACCTTCAGTTTCTCTGAGAGATTCTGAAGCT	2130
OY	1980	aaacaagcctgaaatgagcctttgtgttctgtgggaactgagctctgttctaaatgagac	2039
Db	2131	GAACTCACTGAAGTGGCTTATTTTGTGTGTGTAATTTTGTGATTATTAAGAGGAGG	2190
OY	2040	tggtcatcagaatatttccatctcttgcaaaatataaacacgtttt---ggaataaa	2096
Db	2191	CTGTCAACACGACATTTTCAATATTTTTCAGAGATTATAGGAGATCTTTACCCGAGATGA	2250
OY	2097	tgataaaacacacatlaaagatgatgcagaatatttccatgatatctgcgtcaatttgg	2156
Db	2251	GGATCTCGAAGCAAAAGACAGAAATAGTAACATGTTTCTGTATTTCTAGTCCCTTGG	2310
OY	2157	tgattattgcttgcagttattcatatgcaggagattattttagcgcagagcaggagaaat	2216
Db	2311	AATTATTTCTTTATTTATTCATTTTTCCTCCAGGGCTTTCACATTTTGGCAAAAGCTGGGAGAT	2370
OY	2217	tttaagatgagatlaagacacttgccttcaagcagatgatatatcagatgattgacct	2276
Db	2371	CCCTCACTAAGCGGCTTGATACATATGTTTTCAGATCCATGCTGAGACAGATGTCAGCTG	2430
OY	2277	gtttgatgaaagaagaacagcagcaggagcttgacaacaatatagccaatagatagac	2336
Db	2431	GTTTGATGACCTTAATAAAACACACTGAGACATGTGAACCAACGCTTGGCAATGATGCGGC	2490
OY	2337	acaattcaagaggagcaacagatgcagggcttgcgtcttaacaaaatgaaactaacat	2396
Db	2491	TCAAGTTAAAGGGCTTTAGGTTCCAGGCTTGCTGTCAATTACCCAGAAATATAGCAAACT	2550
OY	2397	gggaacttcagttacatcttcccttataatgatatgtagatgtagatcctgattcttag	2456
Db	2551	TGGGACAGGACATTATTTATCTTAACTAATGAGTTGGCAATTAACACTTTAACTCTTAGC	2610
OY	2457	tattgctccagctacttgcgtgcaggaatgatitgaacccgacgaatgactgatttgc	2516
Db	2611	AATTGTACCATATTCGCAATTAACAGAGTGTGTGAAGAAATGTTGCTGCGCAAGC	2670
OY	2517	caacaagaatlaagaagaacttlaagatgctgtgaagaatgagaactggaagcttgtagaa	2576
Db	2671	ACTGAAGATTAAGAAAGCTTAAGAGAGCTGGGAAGATTGCTTACAGAAACCATCGAATA	2730

OY	2577	tatcgactatagtgctcattaaacgaaggaaaagccttcgcagcaatgcttgtaagat	2636
Db	2731	CTTCGGAACCTGTTCTTTCTTTGACTCGGGGCGCAGAACTTTGAAATCAATGATGACACAG	2790
OY	2637	gcttcagactcaacaacagaataaccctcgaagaagacacagatattctggaagctgtatgc	2696
Db	2791	TTTGCAAGTACCAATACGAAACCTTTGAGGAAGCAGCACTTCGCGGTCTCATTTTC	2850
OY	2697	attcagccatgctcttataattcttgctctatgcagcaggtcttgatctggaactatt	2756
Db	2851	TATCACCCAGCAATGATGATATTTTCTATGCTGCGCTGTTCCGGTTGGTGCTACTT	2910
OY	2757	aatcaagcttggacgaatgagccccaaggagcatgctcatagtttcttaactgtcaatgtca	2816
Db	2911	GGTGGCAAAATGAGTTTCATGAACTTTTCAGAGATGTTCTTTTGGTATTTCTCAGCATATGCTT	2970
OY	2817	tggagcatgcccactcggaaaaacgctcgtcttgctgccttgataatcccaagccaaac	2876
Db	2971	TGGTGCCATGGCACTGGGGGAGGTCACTTATTTGCTCTGACTATGCCAAGCCAAAGT	3030
OY	2877	gagggcttgcgaatcgtcttgccttgcttggaaaaaaagaaaccaaatatagaacgcgagtca	2936
Db	3031	ATCAGACGCCACGCTCATCATGATCATTTGAAAAAACCTCTGATTTAGACGTACAGCC	3090
OY	2937	agaaaggaaaaagccagacacatgtgaaggaaattagaagtttcgaagaagctcttctt	2996
Db	3091	TCACGGGCTCAAGCCAAATCGTTGGAAGGAAATGAGCAATTAATGATGAGTCGTGTTCAA	3150
OY	2997	ctatccatgtccccaatagttttcaatcccgctgctatcccccaglatggaagcag	3056
Db	3151	CTATCCCACTGACACAGACATCCCGCTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAGG	3210
OY	3057	aaagacaglatgacatttcttgaggagcagcgcgtcttgaggaaaagacactctgtcaactct	3116
Db	3211	CCAGACGGTGGCCCTCGTAGTAGTAGTGCGTGTGGGAAGACACAGTTTGACGCTCT	3270
OY	3117	gcgagagacttatagaacccgctgcgaagacaagtgctgttltgagtgctgagtgcaaaaga	3176
Db	3271	AGACGCGTTCTATGACCCCTTGCTGGCTGAGGTTCACTGATTAATGATGGCAAAAGATATAACA	3330
OY	3177	attgaatgtacagtggtctccgttccccaatagaaatcgltccttaaggagcctgtgctct	3236
Db	3331	CTGAATATGTCAGTGGCTCGGACACACCTGGGCACTGTTGTCAAGAGCCATCTGTT	3390
OY	3237	caactgcagcattgtctgaagaacatgcacctatgtgtgaacaacgcgtgtggtgcattaga	3296
Db	3391	TGACTCGAGCATTCGCCGAAGAACTTGGCTCTATGAGACACACCGGGTCTATCATATGA	3450
OY	3297	tgagatcaaaagaagccgcaaatgtcagcaaatatccatcttatttgaaggtctccctga	3356
Db	3451	AGAAATATATGAGGACGACCAGGAGGACCAATACACCTTATCTGAGACACTCTCGTA	3510
OY	3357	gaatatcaaacacacaaagtgtgactggaagaagagaaagccttcttgccgcgcgaanaaag	3416
Db	3511	GAATATCAACACACAGATATGAGACAAAGGAACCCAGCTCTTGTTGGTGCCTCGAAAACAGC	3570
OY	3417	actagctatttgcagaagggcctcttccaaaaaacccaaatattatcttgcgtggaagccac	3476
Db	3571	CATTGCCATATGCTGGCGCTCTTGTTAGACAGCCTCATATTTTGGCTTTTGGATGAAGCTAC	3630
OY	3477	ttcagccctcgataatgtacagtgagaagatggtgtcaagcatgcacctgtgataagccagac	3536
Db	3631	ATCAGCTCTGTGATACAGAAAGTAAAAAGTGTGCCAAGAGCCCTGCGCAAAAGCCAGGA	3690
OY	3537	ggagaaagacatgctctagtgctactcaacggcctctctgaaattcgaacgcagattgat	3596
Db	3691	AGGCGGCAACCTGCACTTGATGATGCCACCGCTTGTCTACCAATCCAGAAATGACAGATTAAAT	3750
OY	3597	agtgctctgcacaaatggaagaataaagaaacaagaaactcatgaagaccccttgcgaaga	3656
Db	3751	AGTGCTGTTTCAGATATGGCAAAATCAAGAGCATGCGACACATCAACAGCTGCTGGCCCA	3810
OY	3657	tcgagacatatatttaagttagtgaatgcacagtcag	3694

```

Db      3811 GAAAGCATCTATTTTCATGTCAGTCCAGGCTG 3848
RESULT  5
LOCUS   AF269224
DEFINITION AF269224 4045 bp DNA SYN 04-JUL-2000
           Synthetic construct Canis familiaris his-tagged-multidrug
           resistance glycoprotein gene, complete cds.
ACCESSION AF269224
VERSION   AF269224.1 GI:8926216
KEYWORDS
SOURCE    .
ORGANISM  synthetic construct.
           artificial construct.
REFERENCE 1 (bases 1 to 4045)
AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE      Analysis of dog MDRI p-glycoprotein
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 4045)
AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
TITLE      Direct Submission
JOURNAL    Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
           Tournefeuille, BP3, Toulouse 31931, France
FEATURES
           source
           1..4045
           /organism="synthetic construct"
           /db_xref="taxon:32630"
           /note="original mRNA isolated from Beagle dog colon"
           101..3964
           /codon_start=1
           /transl_table=11
           /product="his-tagged-multidrug resistance glycoprotein
           MDR1"
           /protein_id="AA81747.1"
           /db_xref="GI:8926217"
           /translation="MHHHHHDPGGRKSGAEKNFMKSKKNEKKKPTVSTFA
           MRRYNNMLRLMYLVTGMAAIIHGALPLMLVGMWTFSEFANAGISRNKTFVINE
           STNNQHFNLHEEMTAYAYYSIGGAGLVAAVYQSFNLAAGROLIKTRKQF
           HAIMROELGMPDHDVGLNTRLTDDVSKINEGIKGFPQSIATPFPFGTVGTR
           GKRLTVIIAISVGLCSAIAKIIISFDKELLYAKAGAAVEVLAIRVIAFG
           GKKELEIRKINLEAKRIGIKAITANISGAFLITVASVLAETVGLSLEY
           SIQVLTVEFSVIGAFSIGQASPIEAFANAGAAVEYFIIDNKPISDSYSGSKH
           PDNKGNLFEKNVHESYPSRKEVKILKGNLQVSOQTVALVNSCGSKSTVOLMR
           LXDPDGMVCIQDODIRITINVRHLREITGVSOEPLFATLIENTRYGSENVYTMDEI
           EKAVENAVYDEIMKLNKEDTLVGERGALSGGKORLAIARVIRNPKIILDEAT
           SALTDESAVYVQALDKARKGRTTIVIAHRLSVNRADVAGDDVYIEKGNHDELM
           KEKGIFFKLVTKOTRGENEIELENATGESSESDALMSPKDSSSLIKRSTKRSIHA
           PGQDRLKLTGKEDLNENPVPSFRLIKLNSTMPYFVGIFCAIINGGLOPSESIIIE
           SRILGIFTRDEDETRONSMNPSVLELVGIIISFTFFLOGTFPGKAGILKRLRY
           MYRSMLRQDVSWFDDPKNTGTALTRLANDAQAQVGAIGSLAVITONIANIGTGI
           ILSIYGMOLFLILLAIPIAIVIAIGVEMKLSGALKDKKELEGACKIATENENT
           VSLTRQKFEFYHQAOSLOVPRYNSLRKAHIEVYSSTIOMMYEYAGCFREGATLV
           ANEFNFQDVLVSAIVFGAMAVGOVSAPDYAAKVSAAHVMITEKSPILDSYS
           PHLKNTLEGNTVNEVFNYPTRPIDPLQSLSEVKGQHLAVGSGCKSTVY
           QLLERYNDPLAGSVLIDGKEIKHLNQWMLAHIGIYQSEIILDCIAEIAVGDMSR
           VVSEHEIMOAKRANIHHEFTLPKYNTKRVGKGLQSGGKORLAIARVIRNPKI
           LILDEATSAIDTESKRVQEAIDKARKGRTTIVIAHRLSTIQNADLIVYQNKVKEH
           GTHQQLIAQKGIFFSMISVOAGAKR"
BASE COUNT 1171 a 850 c 986 g 1038 t
ORIGIN

```

```

Query Match      33.38; Score 1231.2; DB 12; Length 4045;
Best Local Similarity 61.28; Pred. No. 7.2e-247;
Matches 2188; Conservative 0; Mismatches 1258; Indels 132; Gaps 7;

```

```

QY 162 cagcttgaccctgattatgttggaataggtgttcgccttgatttttgatacatca 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 460 CACGTATGCCATATATATACATGCGATCGGTGCTGCTGCTGCTGCTGCTTATCAATCCA 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 222 gatttccttgatataactcagcagcagcagcagcagcagcagcagcagcagcagcagc 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 520 GGTTCATCTCGTGGCCCTGGACAGAGNAGACATACCTCAAAATTAAGAAACAATTTT 579
QY 282 tcaattcagtttggcaccagagacatcgctgttgatagctgtgacatcggtgaactaa 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 580 TCATGCTATCATCGACAGAGATGAGATGGCTGGTTGACGTGATGACGTGGAGCTTAA 639
QY 342 cactcgcattcacagacatg---acaaatcagtgatglatggaataagatgctct 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 640 CACCGGCTCACAGACATGTCTCAAAATCAATGAAGAAATGGCGACAAAATTTGGAAT 699
QY 399 gttgttccaacaatgtctactcttlltgatgtgcctgcgcagtggtgttgtagaggtcg 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 700 GTTCTTTCAATCAATACCAATTTTACACGGTTTATATGAGGTTTACACGTTGTTG 759
QY 459 gaactcaccctagtgactatccacgctccctactataatgcttcaacgagcagatg 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 760 GAAGCTAACCCCTTGATTTTGGCCATCACCCTGTTCTTGACATTCAGCCGCCATCTG 819
QY 519 ttctaagatgltcatctcatctgaaccagtaagaaatgaagtcattccaagctgggc 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 820 GCGAAAGATACTATCTTACTTACTGATTAAGAACTCTTGCTGATGCAAAAGCTGGAGC 879
QY 579 tctgacagaagaagctctgtcatcatcgaacagatcatagcctttagggccagagaa 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 880 AGTAGCTGAAGAAGTCTTACAGCAATCAAGACTGTGATTCCTTTGGAGGAGAAAAGAA 939
QY 639 agacttccaaggtcttccctttaaataacaagatagctgtgttatttccca 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 940 AGAAC----- 944
QY 699 gtgctactaagtgtgtctgtnttltgtaaggtatatacagaatccaaagatgcaa 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 945 -----TTGAAGGTACAAACAAAATTTAAGAGACTAA 978
QY 759 ggaatttgacataaagaactatagcttcaaaagtgctctgtgtgtgtgtgtgtctt 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 979 AAGATTTGGATTAAGAAACCTATCACAGGCCAATTTTATTTGTCGCCCTTCTTATTT 1038
QY 819 taigaatggaaccatagactgtcttlttgatgtaagaacctgtgacttcaatgaga 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1039 GATCTATGCACTATATCTCTGCTTCTGTGTATGGACCTCCTGGTCTCTCCAG--- 1095
QY 879 acctgatacatcgcgagactgtctgtgttcttctttagtgaatccatagcagtta 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1096 ---TGAATATTCTATTTGACAGACTACACTGCTCTTCTTCTGTATTAATGGGCTTT 1152
QY 939 ttgatttgagcagcagctccctcacttctgaaccttgcgaatccgaggaagctgctt 998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1153 TAGATTATTTGACAGGCAATCCCAAGCATTTGCAATTTGCAAGCGAAGGACAGCTTA 1212
QY 999 tcatatttccagttatctataagaaccagatatagaacttcccaagctgata 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1213 TGAATTTTCAAGATTAATTGCAATTAACAGCATTTGACAGCTATGCAAGAGTGGACA 1272
QY 1059 taaactgaaatccatagaagaagctgtggaatttaaaatgttcttccaatcatc 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1273 TAAACCGAATAATTAAGGAATTTGCAATTAAGGAATTTGCAATTAAGGAATTTGCAAT 1332
QY 1119 aagacatctatacaagcttctgaaggtctgaatctcaagaatgaagctgtgagaagcagt 1178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1333 TCGAAAGAAAGATTAAAGATCTTAAAGGCTTCAACCTGCAAGCTTCAAGTGGAGAGAT 1392
QY 1179 cgcttggtcgtctcaatagcagtggaagagatgagtagcagcttcgagaggt 1238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1393 GGGCGCTGTTGGAAACGTGGCTGCGGGAAGAGCACACCGTGCAGCTGATGCAAGAGCT 1452
QY 1239 atatgacccgatgatgcttcatcagctgagctgagatgaatgaatgaatgaatgaat 1298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1453 CTATGACCCCAAGATGATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1512
QY 1299 gggcattatcagaacatatttgagatgtgttagtaagaagcgtgtgtgtgtgtgtgtgt 1358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1513 AAGCATCTTCGGAATAATTACTGCTGTGTGTGATGATGATGATGATGATGATGATGATGAT 1572

```

QY	1359	calccaglaacaaalcalcaagla	tgagacgagatga	tgctgac	gtgaagaagat	tgagagagc	1418
Db	1573	GATAGCTGAAACAAATTTCGCTATG	GGCGCGAAATATGTCCACATG	ATGATGAGTATGAAAGC	1632		
QY	1419	agcaaggaagcaaatg	cgatgatattatca	tggaagttcc	taataatt	taatacaat	1478
Db	1613	TGTTAAAGGAAGCCAAATG	CCATATATTTTTCATGGAAC	TACTAATTAATTTTGAAC	TCT	1692	
QY	1479	ggtatgggaaaaagagat	ctaaat	ttgag	tgagagagcaaa	agagatcgcaat	1538
Db	1693	GCTTGGAGACAGAGGGGGCC	AGCTGAGTG	TGGACAGAAACAGAAAT	CGCCATTGCTCG	1752	
QY	1539	tgccctagttcgaaacccc	caagaattcga	ttttaaga	tgaggtacg	tctgccc	1598
Db	1753	GGCCCTGTTGGCAACCCCA	AGATCTTCTGCTGCATG	AGCGCAACGTCGACTG	GAAC	1812	
QY	1599	agaaagcaagtcagctg	ttcaagctg	cactggaagaagc	gagcaaggtc	gagatacaat	1658
Db	1813	TGAAGAAGGAAGCAGTGG	TTCAGTGGCCCTGGATTA	AGGCCAATAAAGCCG	CGGATACAT	1872	
QY	1659	cggtgatgacacagact	tctcactc	tctgaag	tgcaattg	atgtatgagccca	1718
Db	1873	TGTATATGCTCATCGTTT	GTGCTACAGTTCTGTA	TGCCGATGTCAT	GTCTGTTTGA	1932	
QY	1719	tggaatgctggcagga	aaagagacatg	ctgaactaa	tgagcaaaac	gaggtlctata	1778
Db	1933	TGAGTCATTGTGTGGAA	AGGAATTCATGATAC	TCAATGAAGAAGAGG	GCATTTC	1992	
QY	1779	ttcaactgtgatg	tcaagagata	ttaaaagc	gtga	-acaga	1837
Db	1993	CAAACTGTGCATGCTCAG	CACAGCAAGGAATGAA	TTTGAGTTAGAAAT	TGCCATCGTGA	2052	
QY	1838	attcactgaagaaga	agaccacactc	ctccctg	gac	ctgtgaag	1889
Db	2053	ATCCAAAGGAAAGTGAAT	GTGCTTGGAAATGTCTCC	CAAAAGATTCAGGG	CTCCAGTTTAAT	2112	
QY	1890	caagtcagactcat	tgacaag	ctgaag	atccaccca	cat	1947
Db	2113	AAAAAGAAAGATCAACT	CGCAGAGATTA	TACATG	CACACAGGCC	CAAGACAAAGCTTGG	2172
QY	1948	-----	-----	cc	tgag	gtc	1979
Db	2173	TACAAAGAGAGACTT	GAAATGAGAAATG	ATCCTC	CAAGTTTCTT	CGAGATTCGAAGCT	2232
QY	1980	aaacaagcc	tgatg	gacctt	tg	tggttctgg	2039
Db	2233	GAACTCAAGTGAATGG	CGCTTATTTTGGTGTG	GTATTTTGTCTATTA	TAAAGGAGG	2292	
QY	2040	tgttcacagag	tatttcc	atcattc	gaaatataa	caacatg	2096
Db	2293	CCCTGCACACGACAT	TTTTCATATATTTTCA	AGAGATTA	TAGGAGATCTT	1ACC	2352
QY	2097	tgataaaca	catlaa	agcatga	gcagaat	tattat	2156
Db	2353	GGATCTCGAAACAAAC	AGACAGAAATAGTAA	CAATGTTTTCTG	ATTTCTTA	2412	
QY	2157	tgatattg	cttg	tcag	tattat	ta	2216
Db	2413	AATTAATTTCTTTATTA	ACATTTTCTCC	AGGGCTTCA	ATTTTGGCAAG	2472	
QY	2217	tttaaga	atga	atlaa	gaactg	gccc	2276
Db	2473	CTCAGCTAAGAGCGCT	TCGATACATG	GTTTTCAGATCC	ATGCTGAGACAGAGATG	2532	
QY	2277	gtttcgtga	aaagaa	acagcag	gagagc	ltgaca	2336
Db	2533	GTTTGAATGACCCATA	AAAAACACAC	CTGAGACAT	TGACAAAC	2592	
QY	2337	acaaat	ccaagag	caacag	gttcc	gag	2396
Db	2593	TCAAGTTAAAGGG	CGTATAGTTC	CCAGGCTG	CTGCATATCC	2652	

QY	2377	gggacttcagltatcaatcccttcctttatatagatgagagatgcacatcccgatctcgag	2456
Db	2653	tgggacacggcgcatwttwttatccctttaatctatagttgtggcatttataccttttacttttagc	2712
QY	2457	tattgcctccagttacttgcgtagcaagaatgattgaaacccgagcaaatgactgatttgc	2516
Db	2713	aatgtttaccatcatttcccaatpaccagagattgttgaataaagaaatgttgtcggcaagc	2772
QY	2517	caacaagaataagaacgaacttlaagaatctgtgnaaagatagcaacttgaactttagaag	2576
Db	2773	actgaagaattatgaanaagagcttacaagagcgtgggaagatttgctacagaanaagcattgaaaa	2832
QY	2577	tatacgtactaagtgatcalttaacaagggaaaaagccttcgagcaaatgatalgaagat	2636
Db	2833	ctttccgaactcttwtgtttcttttgactcggggcagaaagtttgaaatcacatgtatgacacagc	2892
QY	2637	gcttcagactcaacaacgaataatccctcgaaagaagacagatattattgaagtgtatgc	2696
Db	2893	ttttgaaagtaccatfaccagaaacacgttttgagaagaaagcacacatcttggaggtccatttttc	2952
QY	2697	attcaagcatgccttataatttcttcctatcgacgaaggttctgatttgaagcctatc	2756
Db	2953	tatcacccagggcattgatgtatatttttttccatagctggcgttttccggtttggggcctactt	3012
QY	2757	aattcaagcttgcagcaatgcagcccgagaaggcaltcatagtttactgcbaatgcata	2816
Db	3013	ggtagcaaatatgatgttcatgtgaacttttcagagatgtttctttgttatcttcagcatattgtctt	3072
QY	2817	tggagctatgycatclogaaaaagcgtcgttttggctcctgtaatttccaagccaatc	2876
Db	3073	tgggtccaatgtgcacatggggcaggtcagtttcattttgcttgcctgactatgccaagccaaaagt	3132
QY	2877	ggggggctgcgcatctgttgcctgtgttgaaaaagaacaaatalagaagccgcagatca	2936
Db	3133	atcacagaccacagctcatctatgatcatattgaaaaaaccccttgatttgacagctacagcc	3192
QY	2937	agaaaggaaaaagccagacacacatgltgaaggaaattagaagtttcgagaagtccttctc	2996
Db	3193	tcacagccttcacgaacaaatagctttggaagaaatgtgacatttaattgaaggtcgtgtccaa	3252
QY	2997	ctatccatgtgtccccaagatgtttcatcctcgttgcctatccctcagtatctgagcgagc	3056
Db	3253	ctatctccacactcagacaaacatcccccggctccacagggcctgcagccttcagatgaacaaagg	3312
QY	3057	aaagaagtagcatcttggggagagcgagcgtgttggaagaacatcttcgttcaacttc	3116
Db	3313	ccagacagctggccctcgttagttagcagtagtgccttgsgaaagacacaaagttttcagctcct	3372
QY	3117	gcagaagcttataagaccgcgtlgaagaagcaagtcgtcttgaatgtgtlgtatgaltgcaaaa	3176
Db	3373	agagagcctttatgacccctttggctgggttcagtgctaatattgattgacaaagagatpaaaca	3432
QY	3177	attgaatgaagtggtctccgttccccaataagaaatcgttctctcaagagcctgtgtcctt	3236
Db	3433	cctgaaatttcagagtggtccgacacacacccggcattcgtgtcaggaagcccatcttcttt	3492
QY	3237	caactgcagacatctctgaagacatgcgcctatgtgtgaacacagccgtgtgtgttccaatga	3296
Db	3493	tgactgcagcatttcccgagaacatttgcttatggagacaaacacccgggtgctatcacatga	3552
QY	3297	tgagatcaaaaagccgcaaatgacgaataatcaatcttcttatttgaagtctccctga	3356
Db	3553	agagatttatgcagaccacccaagagggcamaacatfaccacacttcatcgaagacactccctga	3612
QY	3357	gaaatcaacacacagatgttgacttgaaggagcagaagcttctgcgcgcagaacaaag	3416
Db	3613	gaattatcacacacagacttagagagacaaagaaacccagctctcgtgtggccacagaacacgc	3672
QY	3417	actagctattgcagaagccttctccaanaacccaaatattatgtttgattgagtgagcac	3476
Db	3673	catgtccatttactcgcgcctcttcttgaacagcctcatatttttgccttttgagatgaagctac	3732
QY	3477	ttcagccctcgtataatagatgaagaagtgtgtcagcatgcctcttgaatgaagccaagcac	3536

Db	3733	ATCACCTTGATACAGAAAGTGA	AAAGCTTGTC	CAAGAACGCTCGGACAAACCCAGCA	3792
QY	3537	gggaagacatgcyccatggtgctca	ctcaacaggctctctgcaatc	taagaacgacagattgat	3596
Db	3793	AGGGCCACACGTGCATGTTGATG	CGCCACCGCTGTC	CCACCATCCAGAAATGCAAGATTAAT	3852
QY	3537	agtgtcttcgacaaatggaagaat	aaagaagaacgaatc	atcaagaagctcctcgagaa	3656
Db	3853	AGTGGTGTTCAGAAATGGCAAGT	CAAGACATGCGACACATCAACAG	CTGTGGCCTCA	3912
QY	3657	tcgaacatacatatttaagtagta	ggacagacagtcag	3694	
Db	3913	GAAGGCATCTATTTTCCATGAT	CAATGCTGCTCCAGGCTG	3950	

Query Match	33.2%	Score 1229.6	DB 4	Length 4317
Best Local Similarity	61.1%	Pred. No. 1.3e-246		
Matches 2187	Conservative 0	Mismatches 1259	Indels 132	Gaps 7
Oy	162	caggtgacccctgataatgttgatagatggtctgctgccttgattttgtttacataca	221	
Db	408	CACGATATGCCATATATTTACAGTGGGATGGGTGGTGGCGCTGGTGGGTGCTTTACATCCA	467	
Oy	222	gatttccttgtagaattataactgcaagcaagcaagcaagcaagcaagcaagcaagctttc	281	
Db	468	GGTTTCATTTCGTGGTCCCTGGCAGCAGCAAGAACAGATATCTCAAAATTTAGAAAAACAATTTT	527	
Oy	282	tcattcagttttggcaagagacatcogtctggtttgatagcgcgtgacatcggtgaactaa	341	
Db	528	TCATCTATTCATATGCAACAGCAGAGATTTGGCTGGTTTGACGTGCATGACGTTGGGGACCTTAA	587	
Oy	342	cactgcagatcacagacatctg--acaaatcacgatatgatatgtgagaataagatgtcct	398	
Db	588	CACCGGCGCTCACAGACGATGTCTCCAAATCATATGAAGAAATTTGGCGCAAAAGTTGGAAAT	647	
Oy	399	gttgttccaanaacatgctcacttttctgatgtgctgcgcgcgcgcgcgcgcgcgcgcgcgc	458	
Db	648	GTTCCTTATCAATCAATAGCAACATTTTTCACCGGTTTATAGTGGGCTTTTACACCTGGTTG	707	
Oy	459	gaactgcacctgagagactctatccacgtctccctcttataatgcttcaagagagacatg	518	
Db	708	GAGGTAAACCCCTTGATTTGGCCATCAGCCCTGTTCTTGAGACTTTAGCCGCCATCTG	767	
Oy	519	tctctgagatgycatcatcatatgacacagtaagaatlaagtgctatttccaaagctgggc	578	
Db	768	GGCAAGATACATCTCTTACTTACTGATAAAGAACTCTGGCCCTATGCAAAAGCTGGAGC	827	
Oy	579	tgtgacagaagaagctctgtcatcaatccgaacagtaacatcctttagggcccgagaa	638	
Db	828	AGTACTCAAGAGAGCTTTAGCAGCATCAGAACTGTGATTCCTTTGGAGGACAAAAGAA	887	
Oy	639	agaactcaaaagctcttcccttaataatacaagaatagcttgatttttccc	698	
Db	888	AGAAC-----	892	
Oy	699	gtgctactaagtlgtlctlgltnttltgtaagtatcacagaatctcaagaatgcaaa	758	
Db	893	-----TTGAAGAGTACAAACAAATTTAGAAAGAGCTAA	926	
Oy	759	ggaatttggcatalaaagaagactatagcttcaaaagtgctcttgctgtgctactctt	818	
Db	927	AGAAATTTGGGATTAAGAAAGACTATCACGGCCAACTTTCTATTTGGTGGCGCTTCTTAAT	986	
Oy	819	talgaatgaaactatgtaactgtcttlttgatgaaagcaacctgtacttctaatagaga	878	
Db	987	GATCTATGATATATATGCTTGCTTGCTTCTTGCTGATGGACCTCTGGTCTCTCCAG---	1043	
Oy	879	accttgatatcacacgaggaactgtctgtcgtgttcttcttagtgaatcatalagaa	938	
Db	1044	---TGAATATACTATTGGACAGGTACTCATCTGCTTTCTTTCTGTATTAATTTGGGGCTTT	1100	
Oy	939	ttgcatitggagcagcagctccctcacttcttgaaacctctgcataagccgaagagctgcctt	998	
Db	1101	TAGTATTTGGACAGCGATCCCAAGCATTTTGAAACATTTTCAAAACGCAAGAGACGCTTA	1160	
Oy	999	tcatttttccaggtctatgatataagaaccagatctagataacttttccaaagctgata	1058	
Db	1161	TGAATCTTCAAGATTAATTGACATAAACCAAGCATTTACAGCTATTTGGAAGAGTGGACA	1220	
Oy	1059	taaacctgaatccatagagaagcagctgtgaaattttaaagtlttcttcaattacalc	1118	
Db	1221	TAAACGAGTATATTTAAGGGAAATTTGGAAATTCAAAATGTTCACCTTCAGTTACCTTC	1280	
Oy	1119	aagaccatctatacaagatcttgaagaagctctgaatctcgaatlaagctctgagagacgt	1178	
Db	1281	TCGAAGAGAGTAAAGATCTTTAAAGGCTCTCAACCTGAAGGTCTAGAGCTGGCGACAGAGCT	1340	

Qy	1179	cgcccttgctgcgtctcaatgycgcagctgggaaagtaacgtatgccagctctgcagaagtc	1238
Db	1341	GGCCCTTGTTGGGAACAGTGGCTGCCGGGAAGACAGACCGCTGCAGTGCAGAGGCT	1400
Qy	1239	ataagacccggaatgagtgcttcaatgctgtagatgaatgacacagagctttaagtc	1298
Db	1401	CTATGACCCACACAGATGGCATGGTCTGTATGTATGACACAGGACATTTAGACATTAATGT	1460
Qy	1299	gcgcgcatlalcgagaccatctggaatggctgtagcaagaagccgtcttgctcggagccac	1358
Db	1461	AAGGCACTTTCGGGAATTAATCTGTTGGTGAGTGCAGAGCCTGTGTTGGCCACAC	1520
Qy	1359	catcaatcaacaatatcaatgacatgacagatgactgactgataagaatgagagagc	1418
Db	1521	GATAGCTGAAGAAACATTCGTATGTGCGCCGCAAAATGACCAATGATGATTTGAAAGAC	1580
Qy	1419	agcaagaagaagaatgcgaatgatttcaatgcctgaggttccctaataatlaaacct	1478
Db	1581	TGTTTAAGGAAGCCAAATGCTATATATTTATCATGAAGACTACTAATTAATTTGACACTT	1640
Qy	1479	gtatggggaataaaggagctcaaaatgagtgagggcagaacaagagatcgcgaattgctg	1538
Db	1641	GGTTGGAGAGAGAGGCGCCCGCGCTGAGTGTGGACAGAAACAGAAATGCCATTGCTCG	1700
Qy	1539	tgccctgattcgaaaccccaagaattcgtattttagatgaggtacgctgcgcctgattc	1598
Db	1701	GGCCCTGTTTCGCAACCCCAAGATCTTCTGCTGATGAGCAAGCTGACGCTCGAAC	1760
Qy	1599	agaaagaagtcagctcgcttcaagctgcaactgtagaagcgagcaaaagtcgagatacat	1658
Db	1761	TGAAGAGGAAGCAGTGGTTAGGTGGCCCTCGATTAAGGCCAAGAAAGCCGACATCAT	1820
Qy	1659	cgctgtagcacccgagacttctactcttctgaagtgagaattgattgtagccctaaaga	1718
Db	1821	TGTATAGCTCATACGTGTTTGTCTCTCAAGTTGTTATGCCGATGTCATTCCTGTTTGATGA	1880
Qy	1719	tggaatgctgagcgagaaagagcaaatgctgaactgaactaagcgaagaagagctatla	1778
Db	1881	TGGAGTCATGTGGAGAAAGAAATCATGATGTAACATCATGAAGAGAGGCACTTACTT	1940
Qy	1779	ttcacttgtagtgcacagataattaaaaagctgtga-acagatggaatcaatgacat	1837
Db	1941	CAAACTGTGTACAAATGCAGACAAGAGGAATGAATTTGATTAGAAATGCGCACTGTGA	2000
Qy	1838	attactctgaagaagaagaccactcaactccctctgcaactcgtgaaag-ccat	1889
Db	2001	ATCCAAAGGGAAGTATGATCTCTTGGAAATGTCTCCAAAGATTTAGGGTCCAGTTTAT	2060
Qy	1890	caagtcagactcatctgacaaagctgtaggaatccaccacaatlaaagata-----	1941
Db	2061	AAAAAGAAATCAACTCGCAGAGATATACATGCACACANAGGCCAAGACAGAAACCTGG	2120
Qy	1942	-----agcttcccgagctctcatataaaatttaagtc	1979
Db	2121	TACAAAAGAGGACTTGAATGAGATATCTTCCTTCAGTTTCTCTGAGAGATCTGCAGACT	2180
Qy	1980	aaacaagccgaatggcccttgctgctcggggaactggcttcgtcttaaatgagac	2039
Db	2181	GAACTCAACTGAAATGGCTATTTTGTGTGTATTTGTGTCTATTAATTAACGGGAG	2240
Qy	2040	tgcttcatcagatatttccatcatcttgcgaataattataacaatgattt--ggaataaa	2096
Db	2241	CGTCGACACGACANTTTTCATATATTTTTCAGAGATTAATAGGATCTTTACCCGAGATGA	2300
Qy	2097	tgataaacaacatlaaagcagatgacagaatltatccatgatatctgcatacttggg	2156
Db	2301	GGAACTCGAAGAACAAAGACAGAAATAGTAACAATGTTTCTGTATTTCTTCTAGTCTTGG	2360
Qy	2157	tgctatttgcttgctgagtaatttcaatgacaggaattatttcgcagagcgagggaat	2216
Db	2361	AATTAATTTCTTTTATTAACANTTTTCTCCAGGGCTTCACATTTGGCAAGAGCTGGGGAAAT	2420
Qy	2217	tttaagagatgagatlaaagacacttgcccttcaagaagcagctgtatatacagatattgctg	2276

Db	2421	CCTCAGCTAAGGGGCTTCGATACATAGTGTTCACATCCATCGCAGACAGAGATCAGTG	2480
Oy	2277	gttcgatgaaagaaagcaacagcacagagagcctgcgaacaacatattagcaatagatatagc	2336
Db	2481	GTTTGATGACCCCTAAANAACACACATGAGACATTGACAAACAGGCTTCCCATGATAGCGGC	2540
Oy	2337	acaaattcaagaggaacaaggttccagatltg'gcgtcttaacacaanaatgcacataacat	2396
Db	2541	TCAAGTTAAAGGGGCTATAGGTTCCAGGCTGCTGTCATTACCCAGAAATATAGCAAACT	2600
Oy	2397	gggacttccgatatcccttccttatatatgatgatgggagatgaacattccgatctcgag	2456
Db	2601	TGGGACAGGCATTATATATCTTAATCTATAGGTTG6CAATTAACCTTTATCTTTAGC	2660
Oy	2457	tattgtccagtaactcgtgcgaagaatgatgttgaacccgagcaatgcattgatcttgc	2516
Db	2661	AATTGTACCCATCTTTGCAATAGCAGSAGTGTGTTAAATGAAAATGTGTCTCGGACAAGC	2720
Oy	2517	caacaaagataagaagaacttaagatcgtcgtgaaagataagcaactgaaacttltgagaa	2576
Db	2721	ACTGAAGATAGAAAGAGCTAGAGAGAGCTGGGAATGTGCTACAGAAAGCATGAGAAA	2780
Oy	2577	tatacgtactatagtgatcaatlaacaaaggaaagacccctcgsgcaaatgatltgaagat	2636
Db	2781	CTTCCGAACCTGTTGTTCTTTTGACTCGGGGAGCAGAAAGTTGAAATACATGATACACAGAG	2840
Oy	2637	gcttcagactcaacaagaataatcccgaaagaagacagatatttgaagtgtgtacg	2696
Db	2841	TTTTCGAAGTACACATACAGAAACTTTTGAGAGAAAGACACATCTTCGGGGTCTCATTTTC	2900
Oy	2697	altcagccatcccttataatttcgcctatgcgaagaggttgcatttggagccattct	2756
Db	2901	TATCACCCAGGCATATGATGATATTTTTCATGCTGGCTGTTCGGTTGCGTCTACTT	2960
Oy	2757	aattcaagcttgcagaaatgacccccaaggagcatgtcatagtcttactgcattgcata	2816
Db	2961	GGTGGCAAAATAGTTCATGTGAACCTTTCAGAGATGTTCTTTTGATTTCTCAGCATTTGCTTT	3020
Oy	2817	tgsgagctatggccaatgcgaaaaagccgcgttggcttcgtatgaatttccaaagccaact	2876
Db	3021	TGTGTCCATGTGGCAATGGGGGAGCTCAGATTCATTTGCTTCCTGACTATGCCAAGGCCAAAT	3080
Oy	2877	ggsggctgcgacatcgttctgcctgttgcgaagaaacaaatatagacagcgcagctca	2936
Db	3081	ATCAGCAGCCACCTCATATGATCATATGAAAACCCCTCGATGAGACAGCTACAGCC	3140
Oy	2937	agaaaggaaanaagcacagacacatgtgaaggaaattagaagttcgaagaatctcttct	2996
Db	3141	TCAGGGGCTTCAGGCAAAATAGGTGGAAGAAATGACATTTAATGAGAGTGTGTTCAA	3200
Oy	2997	ctatccatgttgcgcccaatgtttctatccctcgltgcctatccctcagatgatgaagcgag	3056
Db	3201	CTATTCACATCTGACCAACAACATCCCGGTGCTCCAGGGGCTTCGAGCTTCGAGTGAATAAGG	3260
Oy	3057	aagaacagtagcatcttggggaggaagcgagctgtggaaagaagcattcgttcaactctc	3116
Db	3261	CCAGACGCTGGCCCTCGTAGTAGCAATGGCTGTGGGAAGACACAGTTGTTCACTCTCT	3320
Oy	3117	gcagagacttatalgaccccgctgcagaggaacagtgcttlttgatgtgtgtgaltgcataaaga	3176
Db	3321	AGAGGCGTTTATATGACCCCTTTGGCTGGTTCAGTGCTAATTAATGAGGCAAGAGATTAACGA	3380
Oy	3177	atlgaaatgaatagtgctcgttccccaataatagcaatcgttctcctaagagcgtgtcctt	3236
Db	3381	CCTGAATCTCCAGTGGGCTCGAGACACCTCGGGCATTCGTCTCCAGAGCCCATCTCTTT	3440
Oy	3237	caactgcagacatctgtagacaatcgcctctatgtgcaacaagcgctgtgtgtgcatttga	3296
Db	3441	TGACTGCGACGCTTTCGCAAGAAATATGCTTATGAGACAAACAGCCGGGTGTATCACATGA	3500
Oy	3297	tgagatcaaaagaagccgcaaatgagcaaatatcattcttbatgaaagtctccctga	3356





Db 1221 TAAACGATATATTAAGGAAATTTGGAAATTCAAAAATGTTCACTTCAGTTACCTTTC 1280  
Oy 1119 aagaccatcatcaagatcttgaagaagctgaatctcgaatctgaatctcggagagacgt 1178  
Db 1281 TCGAAAAAGATTTAAAGATCTTAAAGGATCTCAACTGAAAGTTGAGAGGGGACACAGT 1340  
Oy 1179 cgccttgcggtcctcaatcggcagcggagagagatcaggtacgtctcctcagaggtc 1238  
Db 1341 GGGCCTGGTGGGAAACAGTGGGTGGGGAAGAGCAGACCGTGGAGCTGATGATCAAGGCT 1400  
Oy 1239 atatagccgagatgcgcttatacagtggtgagatgaagatgacatcagaagcttcaatgt 1298  
Db 1401 CTATACCCCAACGATGATGCTGTATGATGATGACAGACCATTAAGACCATTAATGT 1460  
Oy 1299 gcgagatcatcagagacatctggagtggttgaatcagaagcctgttgcctggagacac 1358  
Db 1461 AAGGATCTTGGGAAATTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1520  
Oy 1359 catcagtaacaaatcaatgaatgaagagatgagatgagatgagatgagatgagatgag 1418  
Db 1521 GATACCTGAAAAATTCCTGATGCGCGGAAATGTCAACATGAGATGAGATGAGAAAC 1580  
Oy 1419 agcaaggagagaaatgcgtatgattatcagagatctcctaaatcaatcaatcalt 1478  
Db 1581 TGTAAAGAGCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1640  
Oy 1479 ggtgggggaaagagatcgaatgagtgagtgagtgagtgagtgagtgagtgagtgag 1538  
Db 1641 GGTGGGAAAGAGAGGCG 1700  
Oy 1539 tgccttgcgaaaccccaagatctgaatttgaatgagatgagatgagatgagatgagat 1598  
Db 1701 GGGCCTGGTGGCAACCCCAAGATCTTCTGATGATGATGATGATGATGATGATGATGAT 1760  
Oy 1599 agaaagcagatcagctgttcaagctgcacgcgagagcgagcaagatcgagatcaat 1658  
Db 1761 TGAAGATAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1820  
Oy 1659 gctgtgagacacagcacttctcactcagatgagatgagatgagatgagatgagatgag 1718  
Db 1821 TGTGATAGCTATCTTGTGTCTACAGTTGATGATGATGATGATGATGATGATGATGAT 1880  
Oy 1719 tggaaatgcgag 1778  
Db 1881 TGGAGTCAATTTGGAGAAAGAAATCATGATGATGATGATGATGATGATGATGATGAT 1940  
Oy 1779 ttcacttgcagatgcagagatataaaagcgtatga-acagatgagatcaatgacat 1837  
Db 1941 CAATCTTGTCAATGACAGACAAAGAGAAATGAAATGAGTAAATGAGTAAATGAGTAA 2000  
Oy 1838 attctacgaaagaaagacaaactcacttccctcagatcgttgaagag-----cat 1889  
Db 2001 ATCCAAAGGTGAATGATGCTTGGAAATGTTCCAAAGATTCAGGGTCCAGTTTAAT 2060  
Oy 1890 caagtcagactcagatgcagagcgtgagagatccacccaatcaataagagata----- 1941  
Db 2061 AAAAAGAGATCACTCGCAGAGATATACATGACACACAGCCAAAGACAGAAAGCTTGG 2120  
Oy 1942 -----agtcctcgtgaagtcctcctataaaatcttaagtc 1979  
Db 2121 TACAAAGAGACTTGAATGAGAAATGTTCTTCACTTCTCTGAGAGATCTCAAGT 2180  
Oy 1980 aaaaagcgtgaagccttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 2039  
Db 2181 GAATCAACATGATGCTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2240  
Oy 2040 tgcctcagatcttccatcactcctcctcctcctcctcctcctcctcctcctcctcct 2096  
Db 2241 CCTGCAACGAGATTTTCAATTAATTTTCAAGATTTATAGGATCTTAACCGAGATGA 2300  
Oy 2097 tgaataaacacataaagcagatgcagaaatctatccatgcataatccgtatcttgcgt 2156  
Db 2301 GGAATCTGAACAAAGACAGAAATGTAATGTTTCTGTATGTTTCTGTATGTTCTGTG 2360  
Oy 2157 tgcatacttgcagatcttcatcagagatcttcaagcagagcagagcagagcagagcag 2216  
Db 2361 AATTAATTTCTTATTAATTAATTTTCTTCCAGGCTTTCACATTTGGCAAAAGCTGGGAGAT 2420  
Oy 2217 ttaacagatgaatlaagacacttgccttcaaaagcagatgataatcagatcttgcgt 2276  
Db 2421 CCTCACTAAGCGGCTTCGATACATGAGTTTTCAGATCAGTGTGAGACAGATGTCAGCTG 2480  
Oy 2277 gttcgtgaaagaaagacagcagagagcgttgcagacaaatataagcagatgataagc 2336  
Db 2481 GTTATATGACCTTAAAC 2540  
Oy 2337 acaaatlaagagagaaagcagatgcagatgcagatgcagatgcagatgcagatgcagat 2396  
Db 2541 TCAAGTTAAAGGGCTATAGGTTCCAGGCTTGTCTGCTATTCACGAAATAGCAATAT 2600  
Oy 2397 gggacttcaagatcacttcccttataatgagatgagatgagatgagatgagatgagat 2456  
Db 2601 TGGGACAGCATTAATTAATTCCTTAATCATGATGATGATGATGATGATGATGATGAT 2660  
Oy 2457 tatgcctcagatctgcgtgcagagatgatgaacgcagacagatgcagatgcagatgc 2516  
Db 2661 AATTTACCCATCATTTGCAATAGCAGAGATTTGTAATGAAATGAAATGTTGCTGCAAC 2720  
Oy 2517 caacaagatgaagcagacacttaagcagatgcagatgcagatgcagatgcagatgcagat 2576  
Db 2721 ACTGAAAGTAAAGAAAGAGCTAGAAAGAGCTGGAAAGATGCTACAGAAAGCCATGAA 2780  
Oy 2577 tatacgaataagatgcatacagaagaaagccttcagacaaatgataagatgagatgagat 2636  
Db 2781 CTTCGAGATGTTGTTTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2840  
Oy 2637 gctcagacacacagaaataactcgaagaagagacagatctatgagagcgtatgc 2696  
Db 2841 TTTGCAAGTACATACAGAACTTTTGAAGAAACACACACATCTTGGGGTCTCATTTTC 2900  
Oy 2697 attcagctgccttataatcttgcagatgcagagatgcagatgcagatgcagatgcagat 2756  
Db 2901 TATCACCCAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2960  
Oy 2757 aatcagctgcagaaatgcagacccagagagcagatgcagatgcagatgcagatgcagat 2816  
Db 2961 GGTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3020  
Oy 2817 tggagctatgcagatgcagaaagcgtcgttgcgtcctgaatctcagagcagaaatc 2876  
Db 3021 TGGTGCATGAGCAGTGGGAGAGTCAATTTGCTCTGATGATGATGATGATGATGATGAT 3080  
Oy 2877 ggggagctgcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagatgc 2936  
Db 3081 ATCAGCAGCCCGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 3140  
Oy 2937 agaagggaaagaaagcagacacatgcagagagatgaagatgcagagatgcagatgcagat 2996  
Db 3141 TCACGCGCTCAAGCAAAATAGCTTGAAGAAATGATGATGATGATGATGATGATGATGAT 3200  
Oy 2997 ctatcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagatgc 3056  
Db 3201 CTATCCCACTCAGCAGAGATCCCGTGTCTCAGGGGCTGAGCCTTCGAGTAAAGAAAGG 3260  
Oy 3057 aagaagcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagat 3116  
Db 3261 CCAAGCGCTGGGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3320  
Oy 3117 gcaagacttataagaccccgatgaagagagatgcagatgcagatgcagatgcagatgcagat 3176  
Db 3321 AGAGCGCTTATGAGCCCTTGGTGGTTCAGTGTAAATGAGCAAGAAATGATGATGATGAT 3380  
Oy 3177 atgagatgcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagatgcagat 3236  
Db 3381 CCGTGAATGCTCAGTGGCTCGAGACACCTGGGCACTGCTGTCTGAGGAGGACCATCTCTGT 3440





QY	1479	ggttgagggaanaagagbctlaaatggtggtgagggcgagaacaagagatcgcgaattgctgc	158
Db	1933	ggttgagagagagagggcccgacgttgagtggtggggcagaaagcagagagatcgccattggcagc	2052
QY	1539	tgcccttagtcgaaccaccaagatcttgattttaagatagagctagctgtgccggttc	1598
Db	2053	tgcccttggttcgcAACCCCAAGATCCTCCTGCTGATGAGGCCACGTAACCTTGGACAC	2112
QY	1599	agaaagaagtcagctgtttaaagctgcacttggagaagagggacaaaggttcgcnatacat	1658
Db	2113	AGAAAGCGAAGCACTGGTGTCAAGGTGGCTCGATGTAAGCCAGAAAGAGTTCGACACCAT	2172
QY	1659	cgtgtgtgcacaacgcgaacttctctactcttcgaagtgcagaattgattgtgacctaaaga	1718
Db	2173	TGTATATGCTCATGTTTGTCATCACTGCTGTATAGCTGACATCACTGCTGGTTTCATGA	2232
QY	1719	tgagatctgcgcgagaaagagacagatctgaactaaatgsgaanaagaggtctatalt	1778
Db	2233	TGGAGCTATTGTGGAGAAAGGAATATATGATGACTCATGAAGGAAGAAAGGCATTTACTT	2292
QY	1779	ctcacctgtgatgtcagaagataattaaagctgatagcagatgagatcaatgacata	1838
Db	2293	CAAACTGTGTCAATGCAGACAGCAGAGAAATGAAGTGAATTAAGAAATGCAAGTGATGA	2352
QY	1839	ttctca-----ctgaagaagaaccaactcaactccctctgcactcgtg	1881
Db	2353	ATCCAAAGATGAAATATGATGCTTGGAAATATGTCCTCAAAATGATTCMAATGCATGCTAT	2412
QY	1882	aagagca-tca-----agtcagatctcatatgcaagagctcgaga	1919
Db	2413	AGGAAAGATCAACTGCTGAGAGTCCGTGGANTACACAGCCCAAGACAGAAAGCTTAG	2472
QY	1920	atccaccaactctaaagaga-taagctctccctgaagctctccatctataaactttaaagt	1979
Db	2473	TACCAAGAGGCTCTGTGATGAAGATACCTCCAGTTCTTTGAGAGATTAATGAAGCT	2532
QY	1980	aaecaagcctgaaatgagcccttctgtgtctctggagacattgctctgttctaattgaa	2039
Db	2533	AAATTTAACTGAAGGCGCTTATTTTGTGTGTGATTTTGTGCATTATTAATATGAGG	2592
QY	2040	tgctacccagatattctcactcactcttgcgaanaatataaacatgltt---ggaataa	2096
Db	2593	CCTGCACACGACATTTTCCAAATTAATTTTCAAGATTAATAGGGTTTATACAGATTA	2652
QY	2097	tgataaaccacatlaaagcatgatagcagaatatlaatacagatattcgctcaatttgg	2156
Db	2653	TGATCTCGAAMCAAAAGCAGACAGATAGTAAGTACTGTTTCACATATGTTTTCAGCCCTG	2712
QY	2157	tgtaattgcttgcagttacttcatgcagagatatttaacgacagaagggaaat	2216
Db	2713	AATTAATTCCTTTATTTACATTTTTCCTTCAAGGTTTTCACATTTGGCAAAAGCGAGAGAT	2772
QY	2217	tttaagatgagatlaaagacacttgcccttcaagaacatgatatatacagatatltgctg	2276
Db	2773	CCTACCAAGGCGCTCCGATACATATGTTTCCATCCATGCTCAGACAGAGATGTAGTTG	2832
QY	2277	gtttgatgaaagaaacacagacagagagctgtgacaacatattatagcatagatatagc	2336
Db	2833	GTTTATGATACCTTAAANAACACACATGAGACATTGACTACCAAGCTCGCCATATGATGCTC	2892
QY	2337	acaaatcgaagagcaacaggtctccagatltggtcgctctaacacaaaatgcaactaacat	2396
Db	2893	TCAAGTTTAAAGGGCTATATGATGCTTCATATGTTTGGCACTACACATGTTACTTTAGC	2952
QY	2397	gggaactctcagttatcttcccttataatgagatgggagatgacatccctgattctgag	2456
Db	2953	TGGGACAGGAATATTTATTCCTTCAATATGTTTGGCACTACACATGTTACTTTAGC	3012
QY	2457	tattgtctcagttactcgtgcagagaatagatctgaaacacgagcaatgacatgattgtgc	2516
Db	3013	AATTGTACCCATCTTGTCAATAGCAGGAGTGTGTTGAATGAANAATGTTGTCTGGCAAGC	3072
QY	2517	caacaaagatagaagaagactlaagatctctggaagaatagacactgaaagcttggagaa	2576

[illegible]

QY	3657	tcgagacataatcattgaagtagtaagtcacagtcacg	3694
Db	4213	GAAGGACATCATTTTTCATGCGTGCAGCTCAGGCTG	4250
RESULT	9		
HUMMDR1			
LOCUS	HUMMDR1	4646 bp	mRNA
DEFINITION	Homo sapiens P-glycoprotein (PGY1) mRNA, complete cds.		PRI 03-DEC-1999
ACCESSION	M14758		
VERSION	M14758.1	GI:187468	
KEYWORDS	P-glycoprotein; drug resistance protein; transport protein.		
SOURCE	Human drug-resistant cell line KB-C2.5 cDNA to mRNA, clones lambda-BHR10, 5, 104).		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 4646) Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B.		
TITLE	Internal duplication and homology with bacterial transport proteins in the mdrl (P-glycoprotein) gene from multidrug-resistant human cells		
JOURNAL	Cell 47 (3), 381-389 (1986)		
MEDLINE	87028230		
REFERENCE	2 (sites)		
AUTHORS	Ueda,K., Clark,D.P., Chen,C.J., Roninson,I.B., Gottesman,M.M. and Pastan,I.		
TITLE	The human multidrug resistance (mdr1) gene. cDNA cloning and transcription initiation		
JOURNAL	J. Biol. Chem. 262 (2), 505-508 (1987)		
MEDLINE	87109132		
REFERENCE	3 (bases 971 to 985; 3095 to 3109)		
AUTHORS	Kloke,N., Tsubota,J., Kakehi,Y., Komano,T., Gottesman,M.M., Pastan,I. and Ueda,K.		
TITLE	P-glycoprotein gene (MDR1) cDNA from human adrenal: normal P-glycoprotein carries Gly185 with an altered pattern of multidrug resistance		
JOURNAL	Biochem. Biophys. Res. Commun. 162 (1), 224-231 (1989)		
MEDLINE	89322246		
COMMENT	[2] sites. Draft entry and computer-readable sequence [1] kindly submitted by I.B. Roninson, 13-AUG-1987. The sequence shown is of a cDNA clone initiating at a minor upstream transcription initiation site and containing the major site of transcription initiation. Location/Qualifiers 1..4646 /organism="Homo sapiens" /db_xref="taxon:9606" /map="7q21" 1..4646 /gene="PGY1" /gene="PGY1" 382 /gene="PGY1" /note="q in [1]; a in [2]" 425..4267 /gene="PGY1" /note="P-glycoprotein" /codon_start=1 /protein_id="AAA59575.1" /db_xref="GI:307180" /translation="MDLEGGNGAKKKKPFKLNKSKDKKKKPYSVSMRYSNWDLKLIVAGTGLAIIHGAGLPLMLVGEWTDIPFANGNLEDMSTNSINDTGTGFNMLEDMTRAYVYSGIAGVLAAYIOVSEWCLAAGRIHKIRQFFAIRMROELGWPEHVEDMTRLENTLDVDVSKINVEIGDKIOMFOSMAEFTGFIYGTGKMLIVLAISPVLGLAAVMAKILISPTDELLAVAKAGAAVEEVLAAIRIVIAFGOKKLEERYVNLFEAKRIGRIKKAITANTSIGAATLLIYASVALAFWGTTLVSGEYSIGOVTRYFESVLICAFSGVAGSPSIEAFANRGAAYEFTKIIDNKPISIDSTSKSHKDNKGINIFRNVHRSYPRKREVKILIKGINKLVQSSQVALVNGSGCKSTVQLMRLYDPEGKVSVDGRIITINVEFLREITIGVSEDEPFLVTTAENIRYCRENVTDEIKAVKEA		
FEATURES	Source		
gene	variation		
CDS			

[illegible]

Db 1180 AGTAGCTGAAGAGCTCTGGACAACTTAGAAGCTGATTGATTTGGAGACAAAGAA 1239  
Qy 639 agactctcaaaaggtcttcccttaataataaagaatagtctgttatttcccca 658  
Db 1240 AGAAC----- 1244  
Qy 699 gtggtactaaagtgtctgtcttnttgtaaagtaacagaatctcaaaagtcgaa 758  
Db 1245 -----TTGAAGGTACACMAAAATTTAGAACACTAA 1278  
Qy 759 ggaatttgcaataaaaggactaagctcaaaagtgctctgtgtgtgtgtgtgtgtgt 818  
Db 1279 AAGAAATTGGGATTAAGAAAGCTATACAGCAATTTATGAGTGCTGTTCCTGCT 1338  
Qy 819 tatgaattgaaactatgtgactgtcttctgtgtatggaacctctgtacttaagaga 878  
Db 1339 GATCTGTGATCTTATGCTCTGGCTCTGTGATGGACCACTTGCTCTCAAGGGGA 1398  
Qy 879 acctgataacacatggagactgtctgtctgtcttctgtatcatcacaagta 938  
Db 1399 -----ATATTCTATTGGACAGTACTCACTGATTTCTTTCTGTATTAATGGGCTTT 1452  
Qy 939 ttgcatctggagcagcagctccctcactctgaaacctcgaatagcccgagagctgctt 998  
Db 1453 TAGGTGGACAGGCAATCTCAAGCATTTGAAGCATTTGCAAAATGCAAGAGACACTTA 1512  
Qy 999 tcatatttccaggttattgataaagaaccagtatagataacttctccagcttgata 1058  
Db 1513 TGAATCTTCAGATTAATGATTAATGAAGCAACTATTGACAGCTATTCGAAGAGTGGGA 1572  
Qy 1059 taacctgataccataagaagagactgtgaaattaaagttcttctcaattcatc 1118  
Db 1573 CAACCAAGATTAATTAAGGAAATTTGGAAATTCAGAAATGTTCACTTCACTTCACTTCA 1632  
Qy 1119 aagaccatcatcaagatcttgaaagtgctgaaatctcagaatlaagcttgagagaagct 1178  
Db 1633 TCGAAAGAGATTAAAGATCTTGAAGGCTTGAAGCTGAAGGCTGAGGTGGGACAGGT 1692  
Qy 1179 cgccttggtcgtctcaatggtcagtgagtggaagtagcgtagtccagcttctgagaagct 1238  
Db 1693 GGCCTGTGGTGGAAAGTGGCTGTGGGAAGACCAACAGTCCACTGTGTGAGAGGCT 1752  
Qy 1239 atatgtctcggatgactggtcttatactgtgtgataagaaatgacatcagaagcttatagt 1298  
Db 1753 CTTATGACCCACAGAGAGGAGTGTCACTGTGTGATGACAGAGATTAATGACCAATTAATGT 1812  
Qy 1299 gtcgacatctcgaagacatattgagtggttagtcaagagcctgttctgtctggaagc 1358  
Db 1813 AAGGTTTCAAGGGAATCATTTGCTGTGAGTCAAGAACTGTATTTGTTCCACAC 1872  
Qy 1359 catcagtaacaatatacaagtaagtagcagagatgactgtgactgtatgaagagatgagaagc 1418  
Db 1873 GATAGCTGAAACATTCGCTATGCGGTGAAATGTCACCATGAGATGAGATTAAGAAAGC 1932  
Qy 1419 agcaaggaagcaaatgtcgtatgatttttttttttttttttttttttttttttttttttttt 1478  
Db 1933 TGTCAAGGAAGCCAAATGCTATGACTTTATCATGAAGAACTGCTCATTAATTTAAACCCCT 1992  
Qy 1479 ggttagggaaaaaagagctcaaaatgagtggaagggagaaacagagagatcgaatctgcg 1538  
Db 1993 GGTGGAGAGAGAGAGGGGCCAGTTGAGTGTGGGCAAGAACAGAGATGCCCATTTGACG 2052  
Qy 1539 tgccttagtctgaaccccaagaatctgatttagatgagctagctctgccttgatct 1598  
Db 2053 TGGCTGTGTTCGACACCCCAAGATCTCTCTGTGATGAAGCCACGTCACCTTTGGACAC 2112  
Qy 1599 aagaagaagtagctgtctcaagctcgaactggaagagagagagagagagagagagagagag 1658  
Db 2113 AGAAAG 2172  
Qy 1659 cgtgtgtagcagacagacttctactatctcgaagtgagatcttgatctgagccctaaagga 1718  
Db 2173 TGTGATAGCTCATCTGTTGTCTACAGTTGCTATGCTGAGAGCTCATGCTGTTTGCATGA 2232

Qy 1719 tggaaatgtcgtgaggaagagagacatgctgnaactaatgagcaaaagagctcatatta 1778  
Db 2233 TGGAGTCATTTGTGAGAAAGGAATCATGATGATCATGAAGAAAGGCAATTTACTT 2292  
Qy 1779 ttcaactgtgatgtcacaagatattaaaagctgtatgaacagatgagtcacata 1838  
Db 2293 CAACCTTGTCAATGACAGACAGCAAGAAATGAAGTTGATTAAGAAATGCACTGATGA 2352  
Qy 1839 ttcta-----ctgaagaagaaccaactacttctctgacactgtg 1881  
Db 2353 ATCCAAAGAGAAATTTGATGCTTGAATGCTTCAAAATGATTAAGATCCAGTCTAAT 2412  
Qy 1882 aagaagatca-----agtaacttactgacagagctgagga 1919  
Db 2413 AAGAAAAAGTCACTGTAGAGAGTCCGTGATCAACAGCCCAAGACAGAAAGCTTAA 2472  
Qy 1920 atccaccaatctaaagagataagctctctctgaagctctctctctctctctctctctctct 1979  
Db 2473 TACCAAGAGGCTCTGATGAAAGTATACCTCCAGTTTCTTTGAGAGATTTGAAGCT 2532  
Qy 1980 aacaagcctgaatggtccttctgtgtctctgtggaatggtcctctgcttaaggaac 2039  
Db 2533 AATTTAACTGATGAGGCTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2592  
Qy 2040 tgttcaatcaagatcttctcaatctctctgcaaaatataacatgtt---ggaataa 2096  
Db 2593 CCTGCAACCGCATTTGCAATTAATTTTCAAGATTAAGGGGTTTCAAGAAATGA 2652  
Qy 2097 tgaataaacacataaagatgacagaatctatcatatcatatctctctctctctctctctct 2156  
Db 2653 TGAATCTGAAACAAAG 2712  
Qy 2157 tgtatttgcctgtcagttatctcaagagagatatttaccgcaagagagagagagagag 2216  
Db 2713 AATTTATTTCTTTTATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2772  
Qy 2217 tttaagatgagattgaagacactgtgctctcaagagagagagagagagagagagagagag 2276  
Db 2773 CTTCCACAGAGGCTCCGATGATGATGTTTCCGATTCATCTCAGACAGAGATGAGTGTG 2832  
Qy 2277 gtttgaatgaagaaagaaacagcagagagagcttgacacacataatgaagcctaaatagc 2336  
Db 2833 GTTTGATGACCCRAAAAGACCACTGAGGATTTGACATCAGAGCTGCCAATGATGTGC 2892  
Qy 2337 acaaatcagaagcaagagttccaggtctgagcgtcttaacacaaatgcaactaacat 2396  
Db 2893 TCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGTGCTGATTAACCAAGATATACCAATCT 2952  
Qy 2397 ggtgaacttcagtaatcattctcttataatgagtagagagagagagagagagagagagag 2456  
Db 2953 TGGGACAGAGATTAATTAATATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 3012  
Qy 2457 tattgtccagtaactgtcgtagaagaaatgactgaagacagcagcaatgagctgattgtc 2516  
Db 3013 AATTTGATCCATATTTGCAATAGCAGAGATTTTGAATTAATTAATGTTGCTGACAGC 3072  
Qy 2517 caacaaagataaagcaagacttaagcagctgtggaagagtagcaactgaaagcttggagaa 2576  
Db 3073 ACTGAAGAGATTAAGAAAGAACTAGAAAGGTCTGGAGAGATGCTACTGAACCAATAGAAA 3132  
Qy 2577 tataagtaataatgtctcaatcaagagagagagagagagagagagagagagagagagag 2636  
Db 3133 CTTCGGAACGGTGTCTTTGACTCAGAGAGAGAACTTGAACATTAATGATGATCTCAGC 3192  
Qy 2637 gttcgaagactcaacaagaaatccctcgaagaagagagagagagagagagagagagagagag 2696  
Db 3193 TTTGCAAGTACATACAGAAATCTTTGAGAGAAAGCAACATCTTTGGAATTTACATTTTC 3252  
Qy 2697 atcagacatgcttctatataatcttgactatgcaagagagagagagagagagagagagag 2756  
Db 3253 CTTACACCAAGCAATGATGATTTTCTTATGCTGATGTTTCCGCTTGGAGAGCTTACTT 3312



OY	999	tcataatttcccgaggtatctgataaagaacaaccagatatagaataacttctccacagctgata	1058
Db	1513	TGAAATCTTCACAAATTAATGATAAATAAGCCAAAGTATTATCCAGCTATTTCGAAGAATGGCCA	1572
OY	1059	taaacctggaatccataagaaggaactcgttgaaattttaaanaatgcttcttcataatccatc	1118
Db	1573	CAAAACCAGATAATATTAAAGGAAATTTGGAATTCCAGAAATGTTTCACTTCACTTACCATC	1632
OY	1119	aagaccatcatcaagattctgaaagtlctgaatctcgaattaaagtcagtcgagagacgt	1178
Db	1633	TCGAAAAGAGTTAAATGATCTTGAAGGGCCCTGAACCTGAAGGTGCAGATGGGCGAGAGGT	1692
OY	1179	cgcccttgctgcgtctaaagcgagcaggggaagagtaacgtaatccaaagtlctcgagaggt	1238
Db	1693	GGCCCTGCTGGTGAACAGAGGCTGTGGGAAGGCACAAACACTCCACGCTGATGAGAGGCT	1752
OY	1239	ataatgctcgaaatgaaatgcttatacatggtatgataagatatacatcaagactttaaagt	1298
Db	1753	CTATGACCCACAGAGGGAGTGGTACGTGTGATGGACAGGATATTAGACCATTAATGT	1812
OY	1299	gcggacatcatcgaaaccataatggagatggttaagttaagaacctgttltgttcggagccac	1358
Db	1813	AAGGTTTCTACGGGAAATCATTTGGTGTGTGATGAGAACCTGTAATTTGGCACACC	1872
OY	1359	catcaatgaacaatcatcaatgaatgaaacgagatgatgtgactatgaagaatgaaagagc	1418
Db	1873	GATAGCTGGAATAATTCGCTGCTATGCCGTGAAATTTCAACATGCATGATGAGATTGGAAGC	1932
OY	1419	agcaaggaaagcaaatcgtglatgatttatacatgaaagtlctccataaatttaatacat	1478
Db	1933	TGTCACAGGAAGCCAAATGCCCTATCTTATCATGTGAACCTGCTCATTAATTTGACACCTT	1992
OY	1479	ggttcgggaaagaaagagccatcaaatgtgtgagaggaagaacaagagatcgcgaattgctg	1538
Db	1993	GTTTGGAGAAAGAGGGGCCCAAGTTGAGTGTGGCGAGAGCAAGAGATGGCATTTGCAGC	2052
OY	1539	tgcccttgctcgaaaccccaagaatltctgaattttaagaatgaagctacgctctgcctgata	1598
Db	2053	TGCCCTGTTGCGAACCCCAAGATCTCTCTGCTGGATGAGGCGACAGTCAGCCTTGGACAC	2112
OY	1599	agaagaacgaatcagctcgttccaagctcgcactcgtgaagaagcgagcaaaagtcgactacaat	1658
Db	2113	AGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGTGATTAAGGCCAAGAAAGGTCGGACACCAT	2172
OY	1659	cgctgtagcaacacgcgaacttctacatactcgaagtgcgaattgattgctgacccataaga	1718
Db	2173	TGTATATAGCTCATTCGTTTGTCTCAACTTGTCTATATGCTGACGTCATCGCTGGTTTCGATGA	2232
OY	1719	tggaatcgtcgcgagaaagagcacatgcctgaactaatgycaaaaagagctctataat	1778
Db	2233	TGGAGTATTGTGGAGAAAGAAATCATGATCACTCAATGAAGAAAGGACATTTACTT	2292
OY	1779	ctcactgtagatctgaacagataataaaagaatgaataaagaagatgagatgaatgacata	1838
Db	2293	CAAACTTGTCAATATGCACAGCAGCAGGAAATGAAGTTGAATTAGAAATGCAAGCTGATGA	2352
OY	1839	tctcta-----ctgaagaagaagaccaactcaactcctctcgcactcgtg	1881
Db	2353	ATCCAAAGATGAATATGATGCTTGGAAATGCTTCAATGAATTCAGAAATCCAGCTAAT	2412
OY	1882	aagagcatca-----agtcgaactcatctgcaagagctcgga	1919
Db	2413	AAGAAAGATCAACTCTGTAGAGTCTCCGTGATCTACAAAGCCCAAGACAGAAAGCTTGA	2472
OY	1920	atccaccaaatctaaagagaataagctctcctgaagctctcctcatataaattttaagtt	1979
Db	2473	TACCAAGAGGCTCTGATGAAGATTAATCATCTCACTTTCTTTGAGAGATTAATGAAGCT	2532
OY	1980	aaacaagcctgaatgaccttctgtgcttctgggagacatgagctctctgtcttaaatgagac	2039
Db	2533	AAATTTAACTCAAAAGGCCCTTAATTTTCTGTGTGATTTTGTGCAATTAATAAGTGAGC	2592

QY	2040	tgtlcalccagatattttccatcactctttgaaaaatataacatgtttc---ggaatata	2096
Db	2593	CCCTGCACACCATTTTCCAAATAAATTTTTCAAAGATTATATGGGGTTTTCACAGAATTTGA	2652
QY	2097	tgtataaaccacatataagcatgtgacgaattatattccatgatatactgtccattcttgg	2156
Db	2653	TGATCTCGAANAACAAACGACAGAAATAGTAACTTGTTTTCATGTGTTCTAGCCCTTGG	2712
QY	2157	tgtatttgctcttgcagattatttccatgcaggattatttciacgagcagcaggggaaat	2216
Db	2713	AATATATTTCTTTATTTACATTTTTCCTTCAGGGGTTTTCACATTTTGGCAAAAGCTGGAGAGAT	2772
QY	2217	tttaacgatgatataagaacacttgcccttccaaagccatgtttatatacagatatctccg	2276
Db	2773	CCCTCACCAAGGGGCTCGGATACATGGTTTTCGATTCATCTCAGATCCAGACAGATGTGAGTTG	2832
QY	2337	acaattcaagagcaacaagtttccaggatttggcgttcttaccacaanaatgcaactaacat	2396
Db	2833	GTTTGATGACCCCTTAAAAACACCTGTGAGCAATTGACTACCAAGGCTGCCAATATATGCTGC	2892
QY	2397	gggacttccagttatcaattcccttataatataatgataatgagatgacatctccgatctga	2456
Db	2953	TGGGACAGAGAAATTAATATTCCTTCATCTATGCTTTGGCAACTAACCTGTACTCTTACG	3012
QY	2457	tattgtccagttacttgcgtgaaagaaatgataatgaaaccgacagcaatgactgattgc	2516
Db	3013	AATGTATACCATCATGTGAAATAGCAGGAGTGTGTAATGTAATAATGTGTCTGGACAAAGC	3072
QY	2517	caacaagaataagcaagaacttaagcatatgctcgtgaagaatagcaactgaagcttggagaa	2576
Db	3073	ACTGAAATAATAGAAGAACTAGAAAGGCGTGGGAAGTGCATCTAGAACATATAGAAAA	3132
QY	2577	tatacgtactatagtgtcaltataacaagggaanaagccttcgagcaaatglatgaagaat	2636
Db	3133	CTTCCGAAACCGTTGTTCTTTGACTCAGGACCAAAAGTTTGAAACATATGTATGCTCGACAG	3192
QY	2637	gcttcaagctcaacacgaaataacttcogaagaagaacagatatcttggaaagcgttatgc	2696
Db	3193	TTTTGAGGTATACCATTACAAAACTCTTTGAGGAAAGCAACACTCTTGGAAATTAACATTTTC	3252
QY	2697	attcagcattcccttataataatttgcctatgcatacgaagggcttgcatttggagccattc	2756
Db	3253	CTTACCCACGCAATGATGCTATTTTCTATGCTGAGATGTTCCGGTTTGGAGCTACTTT	3312
QY	2757	aattcaagcttgagcaatgacccacagaggagcatgtlcatagtttcttaactgcaattgtcata	2816
Db	3313	GGTGGCACATTAACHTCAGTAGCCTTTGAGAGTGTCTGTAGTATTTTACACTGTGTGCTTT	3372
QY	2817	tggagctatggtccatcggaaaaaagcgtcgttcttgcctccgtgaatattccaaagccaatc	2876
Db	3373	TGGTGCATATGGCGGCGGGAAGTCAAGTTCAATTTTGCTCTGCACTATGCCAAAGCCAAAT	3432
QY	2877	gggggctcgcgacatcgttctgccttcttggaaagaacaaataatagacgcgcgcgtca	2936
Db	3433	ATACAGACCCCATCATCATGATCATTTGAAAAAAACCCCTTGTGATTGCACTACACAC	3492
QY	2937	agaagaaggaaaaagcagacacatgltgaagggaatttagagtttcggaagatctcttctc	2996
Db	3493	GGAAAGCCCTAATGGCGGAACACTTTGGAAGAAATGTCCATTTTGGTGAAGTTGTATTCAA	3552
QY	2997	ctatccatgtgcgccagatgttttcaatccctccggtgtcctatccctcagttatitgagag	3056
Db	3553	CTATATCCACCCGACCGGACATCTCCAGTCTTCAGGAGCTGAGCCCTGGAAGGAAAGGG	3612
QY	3057	aaagacagttagattttggggagcagagcgtctgtggaaagcaactctgttcaacttc	3116
Db	3613	CCAGACGCTGGCTTCTGGTGGGCAAGCATGTGCTGTGGAAGAGCAAGTGTGTCCAGTCTCT	3672
QY	3117	gcagagacttatatgaccccgltgcaaggacaagtgcgtgtttagtgtgtgtgaatgcaaaa	3176



```

Db 3673 GGAGGCGTTCTACGACCCCTTGCGAGGGAAGTGCCTTGATGCGCAAGAAATTAAGCG 3732
Oy 3177 attgaattacagtgcccttcctccaaatagcaatcgcttcctccaaagccgctct 3236
Db 3733 ACTGAATGTTCACTGAGTCCGACACACCTGGCATTCGTCACAGAGCCATCCTGTT 3792
Oy 3237 caactgcagcatctgtgaaacatcgctatgltgacaacagcgltgltgccaatga 3296
Db 3793 TGACTGCAGCATGCTGGAACATTCGCTATGAGCAACACCGGCTGTGTCACAGSA 3852
Oy 3297 tgaagtcacaagaagccgcaaatgcaagcaaatccatctcttatttgaagctccctga 3356
Db 3853 AGAGATTCGTGAGGAGGAGCAAGAGGCCAATACATGCTTCATCGATCAGTCACTGTA 3912
Oy 3357 gaataacacacacacagatgtagcgtgaaagagagacgtctctgctgcgacaaacaaag 3416
Db 3913 TAAATATAGCACTAATAGTAGAGACAAAGAACTACGCTCTGCTGGCCACAAACACAG 3972
Oy 3417 actagcatctgcaagggctctctccaaacaaatatttattgttgtagtagagccac 3476
Db 3973 CATTCGCATAGCTCGTGCCTTGTAGACAGCTCATATTTGCTTTGATGAGAACCCAC 4032
Oy 3477 ttcagccctcgataatgacagtgagaggtggttcagcatcccttgataaagcaagac 3536
Db 4033 GTCACCTCTGATACAGAAAGTAAAGTTGTCAGAGAGCCCTGAGCAAGCCAGAGA 4092
Oy 3537 ggaagagacatgcctgaagtgtagctacagagctctcgtgaattggaagacagattgat 3596
Db 4093 AGCGCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4152
Oy 3597 agtggtctgcacaaatgaaagataaagaaacaaagaaactcaagagctcctgagaa 3656
Db 4153 AGTGTGTTTCAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 4212
Oy 3657 tcgagacatatatttgaagtgtagtgatgacagtcag 3694
Db 4213 GAAAGCATCTATTTTTCATGTCAGTGCAGGCTG 4250

RESULT 11
AR028671 LOCUS AR028671 6505 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 5 from patent US 5858744.
ACCESSION AR028671
VERSION AR028671.1 GI:5940644
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 6505)
AUTHORS Baum,C., Stocking-Harbers,C. and Osterlag,W.
TITLE Retroviral vector hybrids and the use thereof for gene transfer
JOURNAL Patent: US 5858744 A 5 12-JAN-1999;
FEATURES
source
1..6505
BASE COUNT 1719 a 1506 c 1606 g 1674 t
ORIGIN
Query Match 33.1%; Score 1225.2; DB 6; Length 6505;
Best Local Similarity 60.9%; Pred. No. 1.3e-245;
Matches 2178; Conservative 0; Mismatches 1268; Indels 132; Gaps 6;

```

```

Oy 282 tcatcagtttgcacagagacagcgtggttgatagctgtgacatcggtgaactaa 341
Db 2272 TCATGCTATATATGACACAGAGATAGGCTGTTGATGACAGATGTTGGAGCTTAA 2331
Oy 342 caactgcagtagaagacattgac---aaatcagtgatgtgatgagataagattgcct 398
Db 2332 CACCCGACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2391
Oy 399 gtgtttcaaaacatgctcacttcttcagatgctgcgtgagctgtgttgtaagagctg 458
Db 2392 GTTCTTTCAGTCAATGAGCAACTTTTTCAGTGGGTTTATAGAGATTATACAGCGTTG 2451
Oy 459 gaactccacttagtgcactatccagctcctccttataatgcttcaagcgagcatg 518
Db 2452 GAAGCTAACCTTGTGATTTTGGCATCAGTCTGTTTGGCATCAGTCTGCTGTCG 2511
Oy 519 tcttgatagtcacatcctatctgacagtaagaatagtgctctcttccaagcgtgggac 578
Db 2512 GGCAAGATACTATCTTCACTTTCATGATTAAGAACTTTAGGCTATGCAAAAGCTGAGAC 2571
Oy 579 tggcagaagaagctctgtcacaatccgaacagtcacatagccttaaggccagagaa 638
Db 2572 ACTAGCTGAAGAGTCTTGCGACAAATTAGAACTGTGATTCATTTGGAGGCAAAAGAA 2631
Oy 639 agaactlcaaaagctcttccctttaaataaacaagatgctgtgttatttccoca 698
Db 2632 AGAAC----- 2636
Oy 699 gtggtactaagtgtgtctgttntttgtaaggtataacagaatccaaagatgcaaa 758
Db 2637 -----TTGAAAGGTACCAACAAATTTTGAAGAGCTTA 2670
Oy 759 ggaatttgcataaaagaagactatagcttcaaaagctgtctgtgtgctgtactctt 818
Db 2671 AAGATTTGGGATTAAGAAAGCTATTACAGCCAAATTTTATAGGTCGCTTCTGCT 2730
Oy 819 tatgaatgaacctatggaactgtcttcttgtatgtaagaaacctcctgattcctaagaga 878
Db 2731 GATTCATGATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2790
Oy 879 acctgatatccatcgcgagcgtcttctgtctgttcttctttagtgaatccatagcaagta 938
Db 2791 -----ATATTTCTATTTGGAAGATCTACCTGATTTCTTTCTGATTAATTTGGGCTTT 2844
Oy 939 ttcgattgagcagcagctgctccacttgaaccttgcacacccgcgagagcgtcctt 998
Db 2845 TAGTGTGACAGGATCTCCCAAGCATTTGACATTTTGCAAATTCGAAGAGAGAGCTTA 2904
Oy 999 tcaatcttccaggtatgataagaaccagatagataacttccacagctgata 1058
Db 2905 TGAATCTTCAAGATTAATGATTAATGAAGCCAGTATTTGACACTTATTCGAAGAGTGCGCA 2964
Oy 1059 taacctgatatccatagaagaagcgtgtgaatttaaaatgttcttccaatccatc 1118
Db 2965 CAACCCAGATTAATTAATTAAGGAAATTTGAAATTCGAATTCCTGATTAACCCATC 3024
Oy 1119 aagacatctatcaagattctgaaggtctgaaatccaaatgaagctgtgagagacagt 1178
Db 3025 TCGAAAGAGATTAAAGATTTTGAAGGCTGTGAAGCTGAGAGTGCGAGAGTGCGCA 3084
Oy 1179 cgccttgcgtgctcacaatgagcagtgaggaagtagtcggtagtcacagctctgcagaggtt 1238
Db 3085 GGCCCTGCTTGAAGAACAGTGGCTGTGGAGAGAGCAACAGTCCAGAGTCCAGAGGCT 3144
Oy 1239 atatgaccgagatgagcttataatcagtgtagatgaagatgacatcagaagctttaaagt 1298
Db 3145 CTATGACCCCAAGAGGGGAGTGAAGTGTGATGACAGCGATTAATGACCATTAATGT 3204
Oy 1299 ggcgcattatcagaccatagagaggtggttagttaaagagcgtgttttgtagggaccac 1358
Db 3205 AAGGTTTCTACGGGAATATATTGTTGTGTGATGATGAGCAACTGTATTTTGTCCACCCAC 3264
Oy 1359 catcagtaacaatataagtatgacagagatgattgtactgtatgaagagatgagagagac 1418

```



Db 3265 GATTACTGAAAACATTCGCTATGCGCGTGAATAATGTCAACCATGATGATGATTTGAGAAACG 3324  
Qy 1419 agcaaggagagcaaatgcatgattatcatgaggttccctaataatlaatacalt 1478  
Db 3325 TGTCAAGGAGCAACCTATGACTTATCATGAAACGCTCATTAATTTGACACCT 3384  
Qy 1479 ggttaaggagaaaggagctcaatgagtgagggcaaaacagagatgcatgtctg 1538  
Db 3385 GGTGGAGAGAGAGGGGCCAGTTGAGTGTGGGAGCAAGAGATGCGCATTTGCACG 3444  
Qy 1539 tgccttaagtcaaaccccaagaattcgaatttagatgagtcagtcgccttgatcc 1598  
Db 3445 TGCCTGGTTGCCAACCACCAATCTCCTGCTGATGAGGCGCATGACCTTGGACAC 3504  
Qy 1599 agaaagcaagtcagctgttcaagctgacgtgagaaagcgcaaaagtcgagtaacaat 1658  
Db 3505 AGAAAGCGAAGCAGTGTTCAGGTGGCTGTGATTAAGCGCAAGAAAGTCGAGACACCAT 3564  
Qy 1659 cgtgttagacacagcacttctactatctgaagtgcaagtttgatgtgacctaagaa 1718  
Db 3565 TGTGATAGCTCATTCGTTGTCTACAGTTCGTATGCTGACGTATCGCTTGTTCATGA 3624  
Qy 1719 tggaaatgctgaggaagaaagacacatgctgaactaactgacaaacgagtcctatata 1778  
Db 3625 TGGACTCATTTGTGAGAAAGAAATCATGATGATGATGATGATGATGATGATGATGAT 3684  
Qy 1779 ttcactgtgcatgacagagataataaaacgtgagacagatgagtcacatgacata 1838  
Db 3685 CAAACTTGTCAAAATGACAGCAGAGAAATGAAATTAAGAAATGAAATGAAATGAAAT 3744  
Qy 1839 ttccta-----ctgaagaagaacccaactcactcctcctcctcctcctcctcctc 1881  
Db 3745 ATCCAAAAGTAAATGATGCTTGGAAATGCTTCAAAATGATGATGATGATGATGATGAT 3804  
Qy 1882 aagagacatca-----agtcagactcattgacaagcctgagga 1919  
Db 3805 AAGAAAGAGTCACTGTAAGAGATGTCGCTGATGATGATGATGATGATGATGATGATGAT 3864  
Qy 1920 atccacccaacttaagaagataagtcctcctgagtcctcctcctcctcctcctcctcctc 1979  
Db 3865 TACCAAGAGGCTCTGAGGAAATATACCTCAGTTCTTGTGAGAGATGATGAAAGCT 3924  
Qy 1980 aaacaagcctgaatgaccttgggtgctgagacatgagtcgctcgtcctcctcctcctc 2039  
Db 3925 AAATTTAACTGATGAGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3984  
Qy 2040 tgtcatcagatatttccatcatccttgcacaaatlaataacacatgtt---ggaataaa 2096  
Db 3985 CCTGCACACGATTTGCAATAATATTTTCAAAAGATTAAGGGTTTTRACAAGATTTGA 4044  
Qy 2097 tgaataaacacacatlaaagcatgagcaaatlaatacctgataatcgtcacttggg 2156  
Db 4045 TCAATCCTGAACAAACGACAGATAGTAATGTTGTTTCACTATGTTTCTAGCCCTTGG 4104  
Qy 2157 tgttatgttcttgcagttatttcatgagagattatattcaagcagcggggaat 2216  
Db 4105 AATTTATTTTATTAATCAATTTTCTTCAAGGGTTTCAATTTGGGAAAGCTGGAAGAT 4164  
Qy 2217 tttaacgatgagattaaagcacttgcctcaaaagcctgataatcagagataatgcg 2276  
Db 4165 CCTCAACAGGGGCTCCGATACATGCTTTCCGATCAGTCATGACAGACAGATGTGAGTGG 4224  
Qy 2277 gttcatgataaagaaagacagcagaggttgacacacataatgacatagatalagc 2336  
Db 4225 GTTGTGATGACCTTAAACACCACTGACCATTTGACTACAGGCTGCAATGATGCTGC 4284  
Qy 2337 acaaatcaagaggaagcaggttccaagatgagcgtcctaacaacaaatgcaactaat 2396  
Db 4285 TCAAGTTAAAGGGGCTATGAGTTCCAGGCTTCTGTAATTAACCAAAATATGCAAAATCT 4344  
Qy 2397 gggacttcaatcattccttatacatgagtgagagatgacattcctgattctcgag 2456

Db 4345 TGGAGCAGGAAATTAATTAATTCCTTCAATATGATGTTGGCAATPAACACTGTACTTACG 4404  
Qy 2457 tattgtccagtaacttgcgtgacaggaatgataacccgaacatgacttgattgc 2516  
Db 4405 AATTTGATCCCATCATTCATGACATGACAGAGTGTGTTGAAATGAATAATGTTGTGGACAAAC 4464  
Qy 2517 caaanaagtaagcaacttaaga tgcgtgaagaatagcaactgaaagccttggagaa 2576  
Db 4465 ACTGAAAGATTAAGAAAGAACTGAAGAGTGTGGAAGAGTGTGACTGAAAGCAATTAAGAAA 4524  
Qy 2577 taacgtactatagtgta tlaacaaggaaacccctcagacaaatgatatagaagat 2636  
Db 4525 CTTCCGAACCGTTGTTCTTTGACTCAGAGAGCAAGATTGTAACATATGATGCTCAGAG 4584  
Qy 2637 gcttcagactcaacagaaataaccccgaaagaagacagatattggaaagcctgtaatgc 2696  
Db 4585 TTTGCGATGACATACAGAAACCTTTGAGAAAGCAACATCTTTGGAATTAATTAATTTTC 4644  
Qy 2697 attcaagcctgccttataatttgcctatgagcagaggttcgaattggagcctatt 2756  
Db 4645 CTTTACCCAGGCAATGATGATTTTCTTATGCTGATGTTCCGTTTGGAGCTTACTT 4704  
Qy 2757 aattcaagctgagcaatgaccccaaggagcagtgatcatttactgcaattgacata 2816  
Db 4705 GGTGGCACAATAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4764  
Qy 2817 tggagctatgacacacagaaacgctcgttggctcctgataatcacaagcaatc 2876  
Db 4765 TGTGTCGATGAGCGGTGGGCAAGTCAATTCATTCCTGCTGATTAAGCAAAACCAAAAT 4824  
Qy 2877 gggggtgagcactgttgccttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 2936  
Db 4825 ATCAGCAGCCCATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4884  
Qy 2937 agaaaggaaagaccccaacacatgltgaaggaaattgaagtttcgaagagtccttctt 2996  
Db 4885 GGAAGGCTTAATCCGACACATGGAAGAAATGACATTTGGAGATTGATTAATCA 4944  
Qy 2997 ctatcagatgcgcagatglttcaactcctcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 3056  
Db 4945 CTATCCACCCGACCGACATCCAGATGCTTCAGGACATGAGCTGAGAGTGAAGAGG 5004  
Qy 3057 aaagacagtagacattgttgagagcagcagcgttgcgttgcgttgcgttgcgttgcgttgcgttgc 3116  
Db 5005 CCAGACGCTGCTCTGTGGGACAGCTGCTGTGGGAGAGACAGCTGCTCAAGCTCT 5064  
Qy 3117 gcagagacttlaagaccctgcaagagcaagtgctgttgaatgagtgatgacaaaga 3176  
Db 5065 GGAAGCGTTCTACGACCCCTTGGCAGGAAAGTGTGCTGATGGCAAGAAATTAAGCG 5124  
Qy 3177 attgaatgacagtggtcgtcgttcccaatagcaatcgttcccaagagcgttgcctt 3236  
Db 5125 ACTGATGTTCACTGCTCCGACACACCTGGGCAATGTTGCCAGAGAGCCATCTGTT 5184  
Qy 3237 caactgcagatgtgcagaaac 3296  
Db 5185 TGACTGACGATTTGCTGAGACATTTGCTTATGAGACACACAGCGGGGTGCTGACAGGA 5244  
Qy 3297 tgaatacaagaagccgcaatgcaagaaatcaatcattcattatgaaagttccctga 3356  
Db 5245 AAGAGATCGTAGAGGCGAAGAGGCAACATACATATGCTTATGAGATGACTGCTTA 5304  
Qy 3357 gaatacaacaacaagttgagctgaaagagcagccttctgcggtccagaacaaag 3416  
Db 5305 TAAATTTATACACTTAATGAGACAAAGCAATCTGCTTCTGCTGCGCAGCAAAACAG 5364  
Qy 3417 actagactgcaagagccttcccaaaaaccccaaaatlaatttgcgttgaatgagcagc 3476  
Db 5365 CATTTGCATGACTGTCGCTTGTGACAGCTTCATATTTTCTTTTGGATTAAGCCAC 5424  
Qy 3477 ttaagcctclogataaagcagtgtagaagtggttgcagcatgacctgataaagcagagc 3536  
Db 5425 GTCAGCTGTGATACAGAAAGTGAAGAGTTGTCAAGAAAGCCCTGGACAAACCGAGAG 5484

0y	3537	gggaagacatgacctagcagcgtcattccaaagctctctcgtcaattccgaacgcagattgat	3596
Db	5485	AGGCCGCACCTGCATATTGATTTGCTCACCGCGCTGTCCACCAATCCGAATGCGACTTAAAT	5544
0y	3597	atgtgtctgcacatgtgaaagataaagaaacgaagactcatcaagagctccctgaagaa	3656
Db	5545	ACTGCTGTTTCAGAATGTGCACAGTCAAGSAGCATGGCAGCATCAGACGTCTGTGCACA	5604
0y	3657	tcgagacatatatttaagttagttgaatgcacagtcag	3694
Db	5605	GAAGGCATCTTAATTTTTCATATGTGTGATGTCACGCTG	5642
RESULT 12			
LOCUS	AR028672	9318 bp	DNA
DEFINITION	Sequence 6 from patent US 5858744.		PAT
VERSION	AR028672		29-SEP-1999
KEYWORDS	AR028672.1 GI:5940645		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 9318)		
TITLE	Baum, C., Stocking-Harbers, C. and Ostertag, W.		
JOURNAL	Retroviral vector hybrids and the use thereof for gene transfer		
FEATURES	Patent: US 5858744-A 6 12-JAN-1999;		
SOURCE	Location/Qualifiers		
	1..9318		
BASE COUNT	2377 a	2217 c	2364 g
ORIGIN			2358 t
			2 others
Query Match			
Best Local Similarity	60.9%;	Score 1225.2;	DB 6;
Matches 2178; Conservative		Pred. NO. 1.3e-245;	Length 9318;
		0; Mismatches 1268;	Indels 132; Gaps
0y	162	caaggtgaccctgtattatgttggaaatagtggttcctgccttgatttttggttacaataa	221
Db	2111	CAGGTAATGCCATTATTTATACAGTGGATTGGTCTGGGGCTGTGTTGCTTACATTCA	2170
0y	222	gatttcctctgtgattatactacgcagcagcacagaccagaagattcgaacaagtttt	281
Db	2171	GGTTTCATTTTGGTGGCTGGCAGCTGGAGACAAATACCAAAATTAGAAAACAGTTT	2230
0y	282	tcaatcagtttggcacaagagacatcgctggtttgtagtcgttgcacatcgttgaactaa	341
Db	2231	TCATGCTAATAATGCGACAGAGATAGGCTGTTGATGTGCACGATGTTGGGAGACTTAA	2290
0y	342	caatcgcaatgaagaacattgac---aaatcgttgatgttatgttgagataaagattgctt	398
Db	2291	CACCCGACTTACAGAGATGCTCTTAAGATTATATGAAGTTATTTGGTGAACAAATTTGAAT	2350
0y	399	gtttgttcaaaacatctcacttttcgaattggcccgagcagtttggttgaagggctg	458
Db	2351	GTTTCTTTCAGTCAATGTGCACAACTTTTCACTGGGTTTATAGTAGATTTTACACGTGGTG	2410
0y	459	gaaactacccctagtgactctataccaagctctcccttataatggtcttcagcggaacatg	518
Db	2411	GAAGCTAACCCTTGATTTTGGCCATCCAGTCCTGTTCTTGTGACATGTACGCTGCTGTCTG	2470
0y	519	tcttagatggttcattctatcttgcaaccgtagtaagtaattgctctatccaagcttgggac	578
Db	2471	GGGAAAGATACATATCTTCACTTTACGTATTAAGAACCTTTAGCGGTATGACAAAGCTGGAC	2530
0y	579	tggtgcaagaagaagcttcttcatcaatccgaacagcatagccctttagggccgaagaga	638
Db	2531	AGTAGCTGAAGAGGCTTTGGCAGCAATTTAGAACTGTGATTTGCATTTTGGAGGACAAAGAA	2590
0y	639	agaaactcaaaagtccttccctttaaataataacaagatatgcttggtttattatlltccca	698

Db	2591	AGAAC-----	2595
Oy	699	gtgactactaagtgtgtctgttnttltglaagflatacacaaatctcaaaatgcaaa	758
Db	2596	-----TTGAAAGGTAAACAACAATAATTAGAAAGAAAGCTAA	2629
Oy	759	ggattttgacataaaagaagactatagctcacaagtgtctctgtgtctgtactctt	818
Db	2630	AAGAAATTGGGATTAAGAAAGAAACCTTTACACCAAAATTTCTAATAGTGTGCTTCCGCT	2689
Oy	819	tatgaatggaaacctatgtagcttcttttggatgtagaaacctctgtagctttaa	878
Db	2690	GATCTATGCACTTATATCTCTGTGACCTTCTGTGATGGGACACACTTGTCTCTCAAGGGA	2749
Oy	879	acctggataaccaatcgagactgtctgtctgttcttctttagtgaatccaa	938
Db	2750	-----ATATTCTATTGGACAACTACTCACTGATTTCTTTCTGTATTAAATTTGGGGCTTT	2803
Oy	939	ttgcaatggagcagcagtlccctcaacttgaaccttgcacaatgcagagagctgtcct	998
Db	2804	TAGTTGTGGACAGCAGCATCTCCAGCACTTGAAGCATTTTCCAAATGCAAGAGCAGCTTA	2863
Oy	999	tcatatttccaagttatgtatagaagaccagatagatagataacttcttccagcgtgata	1058
Db	2864	TGAAATCTTCAAGTAATTAATTAATTAAGCCAAATATTGCACACTTTTGGAAAGATGGGCA	2923
Oy	1059	taaacctbaatccaaagaagaacttggaaatctaaatgttcttccaatccatc	1118
Db	2924	CAAAOCAGATATATTTAAAGGAAATTTGGAAATTCAGAAATGTTCACCTTCAGTTACCCATC	2983
Oy	1119	aagaccatctacaagttcttgaagaagttcgaatcgaatlaagttcttgaagagcgt	1178
Db	2984	TGCAAAATAGATTAAGATCTTTGAAGGGCCCTGAACCTGAAAGGTGCAGAGAGTGGGCAACAGCT	3043
Oy	1179	cgcccttggctgtctcaatgtagcagttggaaagatlaagtafctccagcttcttgcagaagtt	1238
Db	3044	GGCCCTGGTTGGAAACAGTGGCTGTGTGGGAAGACCAACAAGTCCAGCTGATGACAGAGCT	3103
Oy	1239	atatgatccggaatgatgtcttatacaatggtggatgtagaaatgacatcaagcttaaatgt	1298
Db	3104	CTATTCACCCCAACAGAGGGAAATGGTCAGTGTGATGACAGAAATTAAGACCATTAATAGT	3163
Oy	1299	gcgcatcatctcagaaccatatgtgagtgtagttagtcaagaagccctgttcttgcggagcac	1358
Db	3164	AAGCTTTCTACGGGAAATTCATTTGGTGTGGATCGACATCGAACCTGATTTGTTGGCCACAC	3223
Oy	1359	catcagtaacaatctcaagatagtagacgaagttgtgtgacatgttgaagaagatagtagagcgt	1418
Db	3224	GATAGCTAAAAACATTCGCTTATGGCCCTGTAATAATGTCACCATGATGAGATGAGAAAGC	3283
Oy	1419	aacaagggaagcaaatcgtaatgtatcttatacatgtagagtttccataaattaatatcat	1478
Db	3284	TGTCAAAGGAACCAATGCTATATGACTTTATCATGAAATGAGCCTCATTAATTTGACACACT	3343
Oy	1479	ggtlaagggaanaaaggagactcaaatgtagtggagggcaagaacagagagatccgaattgtctcg	1538
Db	3344	GGTTTGGACAGAGAGGGGCCAGTTGATGCTGTGGGCAAGACAGAGATGCGCATTTGCACG	3403
Oy	1539	tgccttaattgaaaccccccaagatcttgcatttagatggaggtccttgcagcttgcctgtatcc	1598
Db	3404	TGCCCTGGTTGGCAACCCCAAGATCTCTCTGCTGATGAGGCCACGTCAGCCTTGGACAC	3466
Oy	1599	agaagaagcagtcagctgttcaagctgcactgtagaagagcgagcaaaagtcgcatcatacat	1658
Db	3464	AGAAAGCAACAGACAGTGTTCAGTGTGAGGTCTGTGATTAAGGCCAAGAAAGGTGCGGACACCAT	3523
Oy	1659	cgtgttagacacacgacttctactatctgaatgtcgaatgtgaattgtgataccctaaagga	1718
Db	3524	TGTGATACCTCATGCTTTGTGTCTACAGTTCCTAATGCTGACAGTCATCGCTGTTTGATGA	3583
Oy	1719	tggaaatgcttgccggagaagaagagacactgtctgaacataatggcgaacacgggtctatata	1778
Db	3584	TGGAGTCAATTGTGGGAAGAAATCAATGATGAATCTATGAAGAGAAAGGACATTACTT	3643

Qy	1779	tttccttgtagtgcacggagataattaaanaagctgatgacagatgvgggtcaatgacata	1838
Db	3644	CAAACTGTGTCAACANTGCAGACACAGGAATGAAGTTAGAAAATGAGAGCTGATGA	3703
Qy	1839	ttctca-----ctgaagaagaacacacatcctctccctgcactctg	1881
Db	3704	ATCCAAAGAGAAATTGATGCGCTTGGAATGTCTTCMAATGATTCGAATCCAGCTAT	3763
Qy	1882	aagagcatca-----agtcagactcattgacaagctgaa	1919
Db	3764	AAGAAAAGATCAACTCGTAGAGTCCGTGCATCACAGCCCCAAGACAAAGCTTAG	3823
Qy	1920	atccaccaactcaaaagagataagcttcctcgaagctccctcatataaatttaagtt	1979
Db	3824	TACCAAGAGAGGCTCTGTGATGAAGTATACCTCCAGTTTCCTTTGGAGAGATTAGAACT	3883
Qy	1980	aaacaagcctgaaatgggccttttggtctctcgggaacattgctcgtctcctaattgac	2039
Db	3884	AAATTTATCTGAAATGGCTTATTTTGTGTGTGTGATTTTGTCCATTATTAATATGAGG	3943
Qy	2040	tgctcatccagtatcttccatcatctcttgcaaaattataaccatgltt---ggaataa	2096
Db	3944	CTTGCAACCGACATTTGGCATAATATTTTCAAAAGATTATAGGGGTTTTCACAGAAATTGA	4003
Qy	2097	tgataaacaacatleaagagatgacagaattatcaccagatccgtcatcttgg	2156
Db	4004	TGATCCCGAAGACAAACGACAGATAGTAACCTGTTTCACATATGTTTCTAGCCCTTGG	4063
Qy	2157	tgatattgcttgtagtatttcaatcagagatbattttaaggcagagcagggaat	2216
Db	4064	AAATTTATCTTTATTAACATTTTCCCTTCAGGGTTTCACATTTGGCAACGCTGGAGAT	4123
Qy	2217	tttaagaatgagatleaagacacttgccctcaagccatgatatagagbatgctg	2276
Db	4124	CCTCACCAACGGCTCCGATACATGCTTTCCGATCCATGCTCAACACAGATGTGACTTG	4183
Qy	2277	gtttgatgaaagaaacagcacagaggcttgacaacaatatagccatagatatagc	2336
Db	4184	GTTTGAATGACCCATAAAACACCCACTGGACATTTGACTACCAAGCGCTGCCAATGATGTCG	4243
Qy	2337	acaaatccaggagcagaacaggtgctcagggttggcgctaaacaaaatgaacataaat	2396
Db	4244	TCAAGTTAAAGGGGCTATAGGTCTTCAGGGCTCGTGAAATACCAGAAATATGCAAACT	4303
Qy	2397	gggacttcagttacatctcccttatataatgataagagacatccctgattctgag	2456
Db	4304	TGGGACAGGAATTAATATATCCTTCATCTATGTTGGCACTAACACCTTATCTTAGC	4363
Qy	2457	tattgctccagtaactgcgctgacaggaatgatitgaacccgacagcatgaactgttgc	2516
Db	4364	AAATTTGATCCCATATTTGCAATATACAGAGAACTTTTGAAGAAAGAAAGTTGCTCGCAACG	4423
Qy	2517	caacaagaatcaagcagaacttgaagatgcttggaagatagacacttgaaagcttggagaa	2576
Db	4424	ACTGAAGATTAAGAAAGCAATTAAGAAGTGTCTGGGAAGATCGCTACTGAAGCAATGAAAA	4633
Qy	2577	tatacgtaactatagtgctcatiaacaaaggaaaaagccttcgagcaaatgatatgaagat	2636
Db	4484	CTTCCGAACCGTTGTTCTTTGACTCAGAGACGAAGAAGTTTGAACATATGTATGCTCAAG	4543
Qy	2637	gtctcagaactcaacagaaatatccctcgaagaagaagacagatatttggaaagctgttagc	2696
Db	4544	TTTTCAGAGTACATACAGAAACCTTTTGGAGGAAGCACACATCTTTGGAAATTAACATTTTC	4603
Qy	2697	atccaagcagtccttatatatatttgcctatcgaagcagaaggtttgatttggacctattt	2756
Db	4604	CTTCAACCAAGGCAATGATGATTAATTTTCTATATGCTGATGTTTCCGTTTGGAGCTACTT	4663
Qy	2757	aattcaagctgggagaaatgaccccgaggcattgltcatagltttactgcgaattgtcata	2816
Db	4664	GCTGGCACATTAATCTATGTAGCTTTTAAGAGATGTCTGTTAGTAAATTAAGCTGTGCTT	4723

Oy	2817	tggagctatgacgaatcggaaaaaacgtcgttlttgctctccatgaatatlccaaggccaatc	2876
Db	4724	TGGTTCCTATGCGCCGGGGCAGTGCAATTTCATTTCTCTCTGTACTATGCCAAAGCCAAAAAT	4783
Oy	2877	gggggcctcgcaatcgtttgacctgttgtgaaaagaacaaccaatafatagacagccgcatca	2936
Db	4784	ATCACACGCCCATCATCATCATATGATCATTTGAATAAAAACCCTTGATTGACAGCTACAGCAC	4843
Oy	2937	agaaaggaaaaagcccagacacactgtgaaaggaaatttagatttcgaaagtctcttctt	2996
Db	4844	GGAAGGCCTTAATAGCCGCAACACATTTGGAAGAATGTCACATTTGGGAGATTGTATTCCA	4903
Oy	2997	ctatccatgctgccagatagtlttcaatctccgttggtcatatccccatgatgtgagcgag	3056
Db	4904	CTATCTCCACCCGACCCGAGCATCCCAAGTCTTAGGAGACTGAGCCTTGAGAGTCAAAGAAAGG	4963
Oy	3057	aagaacagtagaatltgtgaggagacgagctgtgtgaaaagcactctgttcaacttct	3116
Db	4964	CCAGACGCTGGCTCTGTGTGGGACAGATGCTGTGTGGAAGACACAGTGTCTCAGCTCCT	5023
Oy	3117	gcaaggaactlttagaccocctgtgcaaaggacaagtgtcgtlttga ttgtgtgtaatgcaaaaga	3176
Db	5024	GGAGGGGTTTCTACGACACCCTTGACAGGGGAAGATGCGTCTGATGTGGCAAGAAATAAAGG	5083
Oy	3177	attgaatgtaacagtggtctccgttctcccaaatagcaaatcgltctctcaagagcctgtgctt	3236
Db	5084	ACTGATATGTTCAAGTGGCTCCGAGCCACACTCGGGCATCGTGTCCAGAGACCCATCTCTGTT	5143
Oy	3237	caacgcagagcatgtgttgaagaacatgcctatgtgtgaacaacagccgtgtgtgtccattga	3296
Db	5144	TGACGTGCACATGCTGTGAACAATTTGCCCTATGTGAGAACACACACCGGAGTGTGTACAGGA	5203
Oy	3297	tgaagtcaaaagaagccgcacaatgcaagcaaatatccaatcttattatugaagtlctcccta	3356
Db	5204	AGAGATCTGTGAGGCGCAGCAAGAGGCCCCAATPATATGCTTCATCGATCACTGCTCTAA	5263
Oy	3357	gaatacacacacacaagtgtgacgtgaagaagagcaagcttctgcgcgcgaagaacaag	3416
Db	5264	TAAATATATAGCATTAAGTNGGAGACAAAGAACTCAGCTCTGTGGGCGCACAAACAMCG	5323
Oy	3417	actagcttltgaaaggagctctctccaanaaaccaaaatttatgttgtatgtaggcac	3476
Db	5324	CATTTCCTCATAGTCTCGGCCCTTGTTAGACAGCCTCTATTTTGTCTTTTGATGATGACCAC	5383
Oy	3477	ttcaagccctcgataatgtacagctgtgagaagtgtgtcagacatgccctgtalaagccagag	3536
Db	5384	GTCACACTCTGGATACAGAAAGTGAAGAAAGTTGTCCAGAGAACGCCCTTGACAAAGCCGAGA	5443
Oy	3537	gggaagagacatgctcctagtggttcaatcacagagctctctgcaatlcagaaacgaattgat	3596
Db	5444	AGGCGCGACCTCATGTTGATTTGCTCACGCCCTGTTCACCATTCAGAAATGACAGACTTAAT	5503
Oy	3597	agtgttcttgacaatagtgaagaataaagaacaagaagaaactatcaagatctcctgagaa	3656
Db	5504	ACTGTGTTTTGCAATGGCAGAGTCAAGGAGATGACGACATCAACAGACTCTCTGGCACA	5563
Oy	3657	tcgagacatatattttaagttagttgaatgcacatgag 3694	
Db	5564	GAAGGCATCTATTTTTCATAGTGTCAAGTGTCCAGGCTG 5601	
<b>RESULT 13</b>			
E02326	LOCUS	E02326 4378 bp RNA PAT 29-SEP-1997	
DEFINITION	Multidrug resistance relating gene derived from human normal cells		
ACCESSION	E02326		
VERSION	E02326.1 GI:2170561		
KEYWORDS	JP 1990100680-A/1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 4378)		



Oy	1839	ttcca-----ctgaaagaagaccactactctccctgcactcgtg	1881
Db	2066	ATCCAAAGTAGAATTAGAGCTTTGGAAATGTCTTCAATGATTCAGATCCAGCTTAAT	2125
Oy	1882	aagagcatca-----agtcgaactctcatctgcaagcgtaga	1919
Db	2126	AAGAAAAGATCAACTGTGTAGAGGTCCGTGTGATCACAAGCCACAGAACAGAAAGCTTAG	2185
Oy	1920	atccaccacactaagaagataagtcctccctgaagtcctctatlaaaatttaagtt	1979
Db	2186	TACCAAGAGAGCTCTGGATCAAAAGTATACCTCCAGCTTCCTTTTGAGAGATTATGAAGCT	2245
Oy	1980	aaacaagcctgaatgagcctcttgtagttctggagacattgagctcttgcttcaattgaac	2039
Db	2246	AAATTTAACTGAATGGCTTAATTTGTGTGTGTGATTTGTGCATTAATAATGAGCG	2305
Oy	2040	tgatccacaglatcttccatcatctcttgcaaaatlataacacatgctt--ggaataa	2096
Db	2306	CCCTGACACAGCATTTGCAATAATATTTTCAAAAGATTATAGGGGTTTTCACAAAGATTGA	2365
Oy	2097	tgataaaacacacttaagaatgatgacagaattatccatgatctgcattctggg	2156
Db	2366	TGATCTCGAAACAAACGACAGAAATGTAACTGTGTTTCACATTTGTTTTCAGCCCTTGG	2425
Oy	2157	tgattatgctcttgcaagtatcttcatctgacaggaattatttcaagcagacagggagaa	2216
Db	2426	AAATATTTCTTTTATATATATTTTCTCTCAGGGTTTCACATTTTGGCAAAAGCGAGAGAT	2485
Oy	2217	tttaagcagatgaattaaagacacttgagccttcaaaagccatgatatcatcagatatgctg	2276
Db	2486	CCCTACCCAAACGGCTCGGATACATGTGGTTTTCGATGCCATCCAGACAGAGATGTGAGTTG	2545
Oy	2277	gttcgaagaaagaaacagcagcagagccttgacaacaatattagacatagatatagc	2336
Db	2546	GTTTGATGACCCCTRAAAACACACCTGGAGCATTTGACTACCAAGGCTCCGCCAATGATGCTGC	2605
Oy	2337	acaaatccaagagcaacaggtctccagatctgagcgtcttaacacaaaatgcaactaacat	2396
Db	2606	TCAAGTTAAAGGGGCTATAGTTCACAGCCTTGCTGTAATTACCCAGAAATATAGCAAACT	2665
Oy	2397	ggagcttcagttatactctcttataatgatgagtagggagatccagatccgattctgag	2456
Db	2666	TGGGACAGGAATATTTATCTTCCTCACTATGAGTTGGCACTAACACAGCTTACTCTTAGC	2725
Oy	2457	tatgctccagtaactgcgtgacaggaatgatgtgaacccgacgaatgactggaatttc	2516
Db	2726	AATTTGATCCCATCTATGCAATAGCAGAGAGTGTGTGAATGAANAATGTTGTCTGCAACAAGC	2785
Oy	2517	caacaaagataagcaagaacttaagcagatcgtgaaagataagcaactgaaagcttggaga	2576
Db	2786	ACTGAAGATTAAGAAACAACTTAAGAGTCTGGGAAGATCGCTACTGAAAGCAATATGANA	2845
Oy	2577	tatagctactatgctgcatctaaacaaggaagaaagccttcgagcaatgatttgaagaa	2636
Db	2846	CTTCGGAAACGTTGTTCTTCTTGACTAGAGACAGAAAGTTTGAAACATATGTATGCTCAGAG	2905
Oy	2637	gcttcagactcaacacagaatacctctgaaagaagcacagatalttggaaagctgtlatgc	2696
Db	2906	TTTTCAGAGTACCATACAGAAACCTTTTGAGGAAGACACATCTTTGGAAATTAACATTTTC	2965
Oy	2697	atlaagcagatgaccttatatatcttgcctatcagacagaggttgcatttggagctatt	2756
Db	2966	CTTACCCAGGCAATGATGATTTTCTTATGCTGAGCTGAGTGTTCGGGTTTGGAGCCTACTT	3025
Oy	2757	aattcaagctggaagaaatgacccacagggacatgctcatagitttctactgcaatgcata	2816
Db	3026	GGTGGCACAATAACTCATGTAGCTTTTGAGAGATGTTTCTTGAATTTTCAGCGTTGTCTT	3085
Oy	2817	tgagagcatatggcactcggaaaaacgctcglttgctccttgataatccaaagccaatc	2876
Db	3086	TGTGTCCATATGGCCGTGGGCAACATCTAGTTCATTTTGTCTGTACTATGGCAAAAGCCAAAT	3145

QY	2877	gggggctcgcgatcctggttcctctggtttgnaaaagaacccaatctatgacgcgcgagca	2936
Db	3146	ATCAGCAGCCCATCATCATCATGTGATAAATAACCCCTTGATTGACGCTACAGCAGC	3205
QY	2937	agaagggnaaaagccagacacatctgtgaaggaaattagagttctcgaagctctctct	2996
Db	3206	GGAAAGCCTTAATGCCGAACACATTGGAGAGAAATGTCACATTTTGGTGAAGTGTATTCA	3265
QY	2997	ctatccatgctgcgccagaatgtttctcatcctcgttgcctctatccctccatgattgaagcga	3056
Db	3266	CTATCCACCCGACCGGACATCCACAGTGGCTTCAGGAGACTGAGCGCTGAGAGTAAGAGG	3325
QY	3057	aagaacagtagcatctgttggggagcagcgcgctgtggaaagaacatctctgtccaactct	3116
Db	3326	CCAGACCGCTGGCTCTGGTGGGACAGCGTGGCTTGGGAAGACACAGTGGTCCACCTCT	3389
QY	3117	gcagagacattatgaccccgctgcaaaagaacaaagtcgttcttggatgggtgtgagatgca	3176
Db	3386	GGAGCGGTTCTTACGACACCCCTTGGCAGGGAAGTGTCTTGATTGGCAAGAATATTAACG	3445
QY	3177	attgaatgtaacagtgctgcgtctcccaaatagaaatcgcttctcctaagaagcgtgtgctt	3236
Db	3446	ACTGAATGTTCAHGGGCTCCGAGACACCTGGGCATCGTGGTCCAGAGACCATCTCTGT	3505
QY	3237	caactgcaagcatctgtctygaacaatcgctctatggttgacaacagcgctgtgtgtccattga	3296
Db	3506	TGACGTTCAGCGATTTGCTCTAGAAACATTTGGCTTATGAGAACACCGCGTGTGTCCACAGA	3565
QY	3297	tgaagatcaagaagccgcgaatcgaaatcaatcaatctcttatttgaagtgctccctga	3356
Db	3566	AGAGATTGTGAGGGCAGCAAAAGAGGCCAACATACATGCTTCATCGAGTCACTGCTTAA	3625
QY	3357	gaatlacacacacacaaglttgactygaagaagagcacacgtcttctggcgccagaaaacaag	3416
Db	3626	TAAATATATGACCTAAAGTATGAGAGCAAAAGAACTCAGCTCTCGGTGGCCAGAAACAG	3685
QY	3417	actgactattgcaaggcgtctctccaaaaacccaaatttatgttgatgtaggcac	3476
Db	3686	CATTGCCATTAACCTGTCCTCCCTGTTAGACAGCCTCATATTTCCTTTGGATGAAGCCAC	3745
QY	3477	ttcagccctcgatataatgaacgttgaagaagtggtgtcagcatgccccttgataaagccagac	3536
Db	3746	GTCAGCTCTGGATACAGAAAGTGAAAGGTTGTCCAAAGAGCCCTGGACAAAGCCAGAGA	3805
QY	3537	gggaagagcaatgacttagtgctcaatcaacagcgtctctgcaatctcaagaacgcagatttat	3596
Db	3806	AGGCCGACCTCGCATTTGTGATTGCTCACCCCGCTGTCCACATCCAGAAATGACGACTTAAT	3865
QY	3597	agtgtgtctgcacaaatgaaagaatlaaagacaaagaactcatcaagaagctcctgtagaaa	3656
Db	3866	AGTGGTGTTCAGAAATGGCAGAGTCAAGAGCAGTGGCAGCAGTACAGCACTCTGTGCACA	3925
QY	3657	tcgagacatatatttaagttagtgatgtagatgcagctcaag	3694
Db	3926	GAAAGGCACTTAATTTTTCATAGTGCAGTGTCCAGGCTG	3963
RESULT 14			
LOCUS	AX012321	8630 bp	DNA
DEFINITION	Sequence 2 from Patent EP0955374.		
ACCESSION	AX012321		PAT
VERSION	AX012321.1	GI:998370	06-SEP-2000
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE	synthetic construct.		
AUTHORS	artificial sequence.		
TITLE	1 (bases 1 to 8630)		
JOURNAL	Baum C.D., Hildinger M. and Osterlag W.P.		
	Retroviral vectors for gene transfer		
	Patent: EP 0955374-A 2 10-NOV-1999;		
FEATURES	HEINRICH PETTE INST (DE)		
	Location/Qualifiers		

```
source 1. .8630
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        /note="Proviral Plasmid-DNA"
misc_feature 1. .160
        /note="Plasmid-R ckgat (pUC)"
misc_feature 1. .8630
        /note="Retroviral Expressionsvektor SfbetaJMSA1"
misc_feature 161. .677
        /note="5'-LTR"
5'UTR 532. .1219
        /note="MSA1 mdrl cDNA"
misc_peptide 1220. .5062
        /note="MSA1"
misc_feature 5215. .5774
        /note="3'-LTR"
misc_feature 5775. .8630
        /note="Plasmid-R ckgat (pUC)"
BASE COUNT 2234 a 2011 c 2189 g 2196 t
ORIGIN
```

```
Query Match 33.1%; Score 1223.6; DB 6; Length 8630;
Best Local Similarity 60.8%; Pred. No. 2.8e-245;
Matches 2177; Conservative 0; Mismatches 1269; Indels 132; Gaps 6;

OY 162 caagttgacctgtatattgtgaaataggctgtcgtccttgatttggtaacata 221
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1555 CAGGTATGCTATATTACAGTGAATTTGGTGGGTGGTGGTGGTGGTGGTGGT 1614

OY 222 gatttcctgtgattataactcagcagcagcagcagcagcagcagcagcagc 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1615 GGTTCATTCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1674

OY 282 tcattcagtttggcagcagcagcagcagcagcagcagcagcagcagcagcagc 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1675 TCATGCTATTAATGCGACAGGAGATAGCTGGTTGATGTGCACAGTGGGG 1734

OY 342 cactcgcattgacagacattgac---aaatcagtgatgattgaaataagattg 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1735 CACCGGACTTACAGATGATGCTCTAAGATTAAATGAAGTTATTGGTGA 1794

OY 399 gtgtttccaataacatgtctacttttccgattgcccgtgcagtggtttgtgaa 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1795 GTTCTTTCAGTCAATGGCAATTTTTCACCTGGGTGTTATATAGATTATAC 1854

OY 459 gaaactacccctagtcctcaccagctcctcctctataataggcttcagcagcag 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1855 GAAGCTAACCCCTTGATTTTGGCCATCAGTCTGTTCTTGAGCTGTCAGCT 1914

OY 519 ttctaggatgltcatctcatgacaglaaagaaatlaagltgctcattccaaagc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1915 GCGAATGATCTATCTTCACTTACTGATTAAGAACTCTTACGTAATGCA 1974

OY 579 tctggcagaagaagctctgtcatcaatccgaacagtcataagcctttagggcc 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1975 AGTAGCTGAAGAGGTCTGGACGAATTAAGAACTGATTCATTTGGAGCA 2034

OY 639 agaactcaagaagcttcctcttaataatacaagaatgcttggtttattttccc 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2035 AGAAC----- 2039

OY 699 gtggactaagtgtgttctgtntttglaaggatatacagaatctcaagaatgcaa 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2040 -----TTGAAGGTACACAAATAATTAGAGAGAGCTAA 2073

OY 759 ggaatttggcataaaagacatagcttcaaaagtgtctctgtgtgtgtgtactt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2074 AAGAAATTGGGATTAAGAAAGCTATTACAGCCAAATTTCTTAGTGCTGCTTC 2133

OY 819 tatgaatggaacctatgactgtcttttggataggaacctcttgatcttaatgaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2134 GATCTATGCACTTATGCTTGCGCTTCTGGTATGGGACACCTTGCTGCTCAGG 2193
```

```
OY 879 acctgatataccatcggagctgttctgtgttcttcttaagtlaatccatagcg 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2194 -----ATATTCTATTGGACAGACTCACTGATCTTCTTGATTAATTAATGG 2247

OY 939 ttgcattggagcagcgctccctcacttgaaccttcgcataatgcccagagctgctt 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2248 TAGGTGTGGACAGCACTCTCAAGCATTTGAAGCATTTTGAAGAGACAGACAGCT 2307

OY 999 tcatatttccaggltatgataagaaccagatataagataactttcccaagctgata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2308 TGAATCTTCAAGATATTATATATAAGCAAGATATGACAGCATTTGCAAGAGTG 2367

OY 1059 taaacctgaatccatagaagaagctgtggaatttaaaatgttcttcaatatac 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2368 CAACACGATTAATTAATTAAGGAAATTTGGAATTCAGAAATTCACCTTCAGT 2427

OY 1119 aagaccctcatataagaatttctgaaggctcgaatctcaagaatgaagctgaga 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2428 TCGAAGAAAGATTAAAGATCTTGAAGGCTGTAACCTTAAGTGCAGAGTGGCA 2487

OY 1179 cgcttgtcgtctcaatgtagcagtgaggaaagtagcgttagccagcttcgagag 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2488 GGCCCTGTTGGAACACGTGGCTGGGAAGAGACCAACAGTCCAGCTGATGCAAGG 2547

OY 1239 atatgacggatgtagtgcttcatcgtgtgtagaagatgacatcagacttaagt 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2548 CTATAGCCCAACAGAGGGGATGTGCTGTGATGGAAGATTAATTAATGT 2607

OY 1299 gcgcatatcagagaccattttagagtggttagtcaagagcgtgttctgtgga 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2608 AAGGTTCTACGGGAAATCATATGGTGTGGTGAAGCAAGACATGATTTGGC 2667

OY 1359 catcagtaacaatalcaatagtagcagagatgtagtgcagatgaagatggaag 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2668 GATAGCTGGAACAACTTGGCTATGGCCGTGAAGATGTCACCATGATGATGA 2727

OY 1419 agcaaggaagcaaatgctgattatcatcagagtttccataaattatcatc 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2728 TGTCAAAGAACCCAACTTCTATCATGAAATCGCTCATTAATTTGACACCT 2787

OY 1479 gtttaggggaaaaagagctaaatgagtgtagaggcagaagaagatgccaattg 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2788 GGTGTGAGAGAGAGGGCCCGATTGAATGTTGGCAAGCAGAGGATCCCATTTG 2847

OY 1539 tgccttagtgaaccaccagacttctgattttagaagagctagctgtccctgata 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2848 TGCCCTGTTGGCAACCCCAAGATCTCTGCTGATGAGGCCACGTCAGCTTG 2907

OY 1599 agaaagcaagtcagctgttcaagctcagcgtgagaagcgagaagatcgagata 1658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2908 AGAAGCGAAGCAACTGGTCTGAGTGGCTTGATTAAGGCCAGAAAGGTGGAC 2967

OY 1659 cgtgtagcacaacgacttttactatttcgaagtgcagatttgaatgtgaccc 1718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2968 TGTATAGCTCATCTGTTGTCTACAGTTCTGTAATGCTGACGTATCTGTTG 3027

OY 1719 tgaatgtcgcggaagaagaagacatgctgaaactaattgcaaaacgagttata 1778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 3028 TGGAGTCAATTGTGAGAAAGAAATCATGATGAATCATGAAGAGAGCAATTACT 3087

OY 1779 ttcaattgtatgacagagatataaaagctgtagaagacagatgagatgacata 1838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 3088 CAACCTTGTCAATTCAGACAGCAGCAAGAAATGAAGTTGAATAATGAGTGA 3147

OY 1839 ttcta-----ctgaagaagaagcacaactcctcctcgtcactctgtg 1881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 3148 ATCCAAAAGTGAAATGATGCTTGGAAATGCTTCAAAATGATTCAGATCAGTGA 3207

OY 1882 aagagcatc-----agtcagactcatgcaagagctgaga 1919
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 3208 AAGAAAAAGATCAACTCTAGAGAGTGTCCGTGATCACAACCCCAAGCAAGAG 3267

OY 1920 atcacccaatcaagaagatagcttctcgtgaagctcctctatlaaaatttaag 1979
```



Db	3268	TACCAAGAGGCGCTGTGATGAAAGTATACCTCCAGTTTCCTTTGGAGGATTTATNAAGCT	3327
Oy	1980	aaacaagcctgaaatgycgcttttggtgtctgcyggaacattgycctctgtctctaaatgagac	20339
Db	3328	AAATTTAACTGAATGGCGCTTATTTTGTGTGGTGATTTTGGCAATTTAAATGAGAG	3387
Oy	2040	tgltcacccaglatlctccatccattcttgcaaaatataaecaatgltt--ggaataaa	20986
Db	3388	CCCTGCACACACATTTTCCAAATAATTTTCCAAAGATTAAGGGGTTTTACAAAGAAATTGA	34479
Oy	2097	tgataaaacccacatlaaagatgatgcagaatatcttccatgatattgcattgcattggg	21566
Db	3448	TGATTCCTGAAACAAACAGACAGATAGTACTTGTTTACTATTTGTTTCTAGCCCTTGG	3507
Oy	2157	tgattattgctctgcatgattatctatctgcagggattattctacgcagacgcyggaat	2216
Db	3508	AAATTATTTCTTTTATTTATCATTTTTTCCCTCAAGTTTCACATTTTGGCAAGCTGGAGAGAT	3567
Oy	2217	cttaacgaltgagatlaagacacttgyccttcaaaagcattatatacgaalatgctg	22767
Db	3568	CCCTCACCAAGGCGCTCCGATACATAGTTTCCATCTCCATGCACAGACAGATGAGTTGG	36227
Oy	2277	gtttgatgaaaaggaacagcagcagggcttgacaaacaatatatagccatgatatagc	2336
Db	3628	GTTTGATGACCTTAAAAACCACTGAGCATTTGATACACAGCTCCGCAATGATGCTGC	3687
Oy	2337	acaatctcaagagcaacagagcttccagagattgycgtcttaacaaaatgcaacat	2396
Db	3688	TCAAGTTTAAAGGCGCTTAGTTCAGAGCTGCTGTAAATTACCCAGAAATATGCAAACT	3747
Oy	2397	ggagacttcaglatcatlctccttatataatgagatgagatgacatcttcattctgag	2456
Db	3748	TGGACACGAGAAATATTTATTCCTTTCATCTGTTGGTGGCAATTAACACTGTACTCTTAGC	3807
Oy	2457	tattgctccagtaetctgctgtgacaggaatgatitgaaccgcagcaatgactggaattgc	2516
Db	3808	AAATGTACCCATCTTTCAATTAACAGAGAGTGTGTAAGAAATGAATTTGTCGGACAAGC	3867
Oy	2517	caacaagatagaagaagacttaagcatgctgtgaaagatagcaacgtgaagcttggagaa	2576
Db	3868	ACTGAAAGATTAAGAAAGAACTAGAGAGTCTGGGAAGATCCCTACTGAAGCAATTAAGAAA	3922
Oy	2577	tatacgtactaagtgatcatctaaagaaggaagacgttcgagcaaatgataagaaagat	2636
Db	3928	CTTCCGAACCGTGTCTTTCTTGAACAGAGAGAGTTGAACATATGTATGCTCACAG	3987
Oy	2637	gcttcagaactcaacagaaataactctcgaaagaagcacagatatttgaagctgtatgc	2696
Db	3988	TTTTCCAGGTACATACAGAAACCTTTGAGGAAGACACATCTTTGGAAATTAACATTTTC	4047
Oy	2697	attcagcagctccttatatatatttgcctatgtagaagaggtttcgatttggagccattt	2756
Db	4048	CTTACCCAGGACATGATGATATTTTTCCTATGCTGATGATGTTTCCGGTTTGGAGCCACTT	4107
Oy	2757	aattcaagcttgaagaaatgacccagagagggcagttcatagtttcttaactgcattgcata	2816
Db	4108	GGTGGACATTAACACTCATGAGCTTTGAGGATGTCTTGTTAGATTTTTCAGCTGTGTCTT	4167
Oy	2817	tgagagctatggcatctggaaaaaacgctcglttggctcttgatattccaaagccaatc	28767
Db	4168	TGTGGCCATGTGCGCTGGGGCAAGTCAGTTTCATTTGCTCTGACTATGCCAAAGCCAAAT	42227
Oy	2877	gggggcttgcgagatcgttttgctgttggaaaaaagaacaatatagagacgcagca	2936
Db	4228	ATCAGCACCCCATCATCATATGATCATTTGAAAAACCCCTTTGATTTGAGAGCTTACAGAC	4287
Oy	2937	agaagggaaaaagccagacacatgltgaaggaatttagagtttcgaaagctctctt	2996
Db	4288	GGAAGGCGCTATGCGCGAACAATTTGGAAGGAATATGTCACATTTTGGTGAAGTTGATTCAA	4344
Oy	2997	ctatcacgtgcgccagatglttctcatccctcgltgcttatccctcaagatattgagcgag	3056

Db	4348	CTH0CCACC	CGGACCGGACATCC	CAATGCGCTTC	CAAGGAGCTAGGCTGAGAGTGAAGAAGG	4407
Qy	3057	aagacagatgac	atctgtgtggtgagcagcggtgtgtggtgaagaacattctgttcaactct	3116		
Db	4408	CCAGACCGCTGGCTCTGGTGGGACACATGGCTGTGGGAAGAGCAGTAGTGTCCAGTCTCT	4467			
Qy	3117	gcagagacttata	gaccccggtgcgaagacaaagtgtctgtttgtatgtgtgtgtatgtgtgaacaaga	3176		
Db	4468	GGAGCGGTTGTACGACCCCTTTGGAGGGAAAGTGCTGTGTAAGGCAAAATAATTAAGCG	4522			
Qy	3177	attgaatgtacagtggtgtccgtgtcccaatlagcaatcgttccctcaagagcctgtgtctt	3236			
Db	4528	ACTGATATTATGATGCGCTCCGAGACACACTGGGCGATCTGTCTCCAGAGCCCATCTGT	4587			
Qy	3237	caactgcagcatctgtctgagaacatctgcctatgtgtgatacaacagcgctgtgtgtccattaga	3296			
Db	4588	TGACTGCGAGCATGTGCTGAAACATTTGCTATGGAGAACAAAGCCGGGTGTGCACAGGA	4644			
Qy	3297	tgaatcaagaagaagcgcgaatctgcagcaaatatccattctttatattgaagtgctccctga	3355			
Db	4648	AGAGATCGTGAAGGCGACGAAAGGAGCGCAACATACATATGCCCTTCATGAGTCACTGCTAA	4707			
Qy	3357	gaataacacacacgaagtgtgacacgtgaagaagacagacgttctctgtgcgcagaaacaag	3411			
Db	4708	TAAATATATGCACTTAATAGTAGGAGACAAAGAACTCACTGCTTCTGTGTGCCAGAAACAG	4765			
Qy	3417	actagctatctcaagggctctctctccaaaacccaataatttattgtgtatgagggcac	3477			
Db	4768	CATTGCCATACCTGTGGTCCCTGTGTATACAGCGCTCATATTTTCTTTGTGATGAAGCCAC	4822			
Qy	3477	ttcagccctcgataatgacagtgaggaagtgtgttcagcatgtcccttgataaagccaagac	3533			
Db	4828	GTCACTCTTGGAATACAGAAATGTAAAGGTTGTTCAAAGAGCCCTGTGACAAAGCCAGAA	4887			
Qy	3537	gggaagagcacagcctagtggtgtcactcaacagctctctgcgaatctcaagaagcgaattgat	3596			
Db	4888	AGGCCGACCTGCATATGTGATGTCTCACCCCTGTCTCCACCATCCAGAAATGCAACTTAAT	4947			
Qy	3597	agtgtcttcgacaatgtaagaataaagaagaacaagaacatctcaagaagctctcgaga	3656			
Db	4948	ACTGCTGTTTCAGAAATGAGCAACAGTCACAAAGAGCATGCGACGCATCAGCACTGTGSCNA	5007			
Qy	3657	tcgagacataatcttgaagttagtgatgtaatgcaggtcag	3694			
Db	5008	GAAAGCATCTATTTTTCATGTGTCATGTCACAGGCTG	5045			
RESULT	15					
AX012320						
LOCUS	AX012320	8630 bp	DNA	PAT	06-SEP-2000	
DEFINITION	Sequence 1 from Patent EP0955374.					
ACCESSION	AX012320					
VERSION	AX012320.1	GI:9998369				
KEYWORDS						
SOURCE						
ORGANISM		synthetic construct.				
REFERENCE		synthetic construct				
AUTHORS		artificial sequence.				
TITLE		1 (bases 1 to 8630)				
JOURNAL		Baum,C.D., Hildinger,M. and Osterberg,W.P.				
FEATURES		Retroviral vectors for gene transfer				
Source		Patent: EP 0955374-A 1 10-NOV-1999;				
		HEINRICH PETTE INST (DE)				
		Location/Qualifiers				
		1..8630				
		/organism="synthetic construct"				
		/db_xref="taxon:32630"				
		/note="Proviral Plasmid-DNA"				
		1..8630				
		/note="Retroviral Expressionsvektor SPbeta7Im4"				
		1..160				
		/note="Plasmid-r ckgat (pUC)"				
		161..677				
		/note="5'-LTR"				



5'UTR	532..1219
mat_peptide	1220..5062
	/note="m4_mdr-1 cDNA"
misc-feature	5215..5774
	/note="3'-UTR"
misc-feature	5775..8630
	/note="Plasmid-R ckgat (pUC)"
BASE COUNT	2235 a 2011 c 2187 g 2197 t
ORIGIN	

QY 2097 tgataaacacattaaagcatgacgagaattatattccatgatattgctcatattgg 2156  
 DB 3448 tcatctctgaaacaaacgacagaaatgtaactgtttctacattgtttctagcccttg 3507  
 QY 2157 tgtatttcttctgtcagttatttccatgcaaggtatttccatgcaagcaggggaaat 2216  
 DB 3508 aattttatt 3567  
 QY 2217 tttaacgatgagatlaagacacttggccttcaagccatgltataltcaagatttgcctg 2276  
 DB 3568 ccttcacacagcgcctccgacatgatgttttccgatctcagacagagatgagtg 3627  
 QY 2277 gtttgaatgaagaaagaaacagcagggaggtctgaacaacatattgcatgatattag 2336  
 DB 3628 gttttgtgaccccttaaaacacacacacacacacacacacacacacacacacacac 3687  
 QY 2337 acaaatcaagagcaacaggttccagatlgcgcgttcaacaacacacacacacacacac 2396  
 DB 3688 tcaagtttaagggccttattggttccagccttgcctgaatttaccacacacacacacac 3747  
 QY 2397 ggaacttcaatcattccttataatgagatggagatgacatccctgattctgag 2456  
 DB 3748 tgggacaggaatattatatttcttactatggttgacaaactgactgttactttagc 3807  
 QY 2457 tattgtccagttacttgcgcgtgacaggaattgaaacccgacacatgactgattgc 2516  
 DB 3808 aattgtaccatcattgacatgacagagttgttgaataaagtgttccgacaaac 3867  
 QY 2517 caacaaga taagcaaacacttaagcatgctgaaagatagcaactgaaacttggagaa 2576  
 DB 3868 actgaagaatgaagaagaactgaactgctggaaagatgcctacactgaacacatataaaa 3927  
 QY 2577 taacgtactatagtgcatlaacaaggaagaaagccttcgagcaaatglatagaagat 2636  
 DB 3928 ctttccgaacccgttgtttctttgactcagagcagaaatttgaaacatattgtctcag 3987  
 QY 2637 gcttcagactcaacaagaataaccccaagaagaacagatattggaagcgttatgc 2696  
 DB 3988 ttttgacgactac 4047  
 QY 2697 attcagcactgacctatataatttgcctatgcaagcaggttccatgattggaagcattt 2756  
 DB 4048 ctttcac 4107  
 QY 2757 aattcaagctgagcaatgaacccagagggcagtgatcatagtttctactgcaattgcat 2816  
 DB 4108 gctgcac 4167  
 QY 2817 tggagctatgacatcgcgaaacacgctcttggctcctgaaatattccaaagcacaac 2876  
 DB 4168 tgggtccatgacgctggggcgaagtcagttcattttgctcctgactatgccaagccaaat 4227  
 QY 2877 ggggagctgacatcgttgccttctgaaagaaacaaatataagacagccgcagta 2936  
 DB 4228 atcagacagccac 4287  
 QY 2937 agaaaggaagaaagcagac 2996  
 DB 4288 ggaagccctaatgac 4347  
 QY 2997 ctatcagtgctgcccagatglttccatcctcgttgctatccctcagtatattgaagcag 3056  
 DB 4348 ctattccac 4407  
 QY 3057 aaagacagtagcatgttgaggagcagcgtctgggaaagacacttctgtaacttct 3116  
 DB 4408 ccagacagctgctctgctgggacacagtgcttgaggaaagacacagtgctccagctcct 4467  
 QY 3117 gcagagacttcatgacccctgcaagacacacacacacacacacacacacacacacacacac 3176  
 DB 4468 ggagcgggttctacgaccccttgacaggaagtgctgtgatgacgaagaaatattgaagc 4527

QY 3177 attgaatgacagtgctccgcttcccaaatagaacatgcttccctcaagacgctgtcctt 3236  
 DB 4528 actgaatgttccagtgctccgac 4587  
 QY 3237 caactgacacatgctgagaaac 3296  
 DB 4588 tgcactcagacattgctgagaaac 4647  
 QY 3297 tgagatcaagaagccgcaatgacgaataatccatcttcttatttgaaggtccctga 3356  
 DB 4648 agagatcgtgagggcagcaagacacacacacacacacacacacacacacacacacacac 4707  
 QY 3357 gaatatcaac 3416  
 DB 4708 ttaatttaccacttaaaagtgaagacacacacacacacacacacacacacacacacacacac 4767  
 QY 3417 actagcttgcagagccttctcccaaaaacacacacacacacacacacacacacacacacac 3476  
 DB 4768 catgacatgacgtcgtcccttctttagacacacacacacacacacacacacacacacacacac 4827  
 QY 3477 ttcagccctcgataatgacagtgagaggtggttcaagatgccttgataaagcagac 3536  
 DB 4828 gtcagctctgagatgacagaaagtgaaaggtgtccaaagaaagcccttgacacaaagccag 4887  
 QY 3537 gggaaagacatgctagtggtcactcagagcctcctgcaatcgaagcgaagattgat 3596  
 DB 4888 agccgcac 4947  
 QY 3597 agtgttctgcac 3656  
 DB 4948 agtgggttgcac 5007  
 QY 3657 tgcagacatataatttaagttgaatgacagtcag 3694  
 DB 5008 gaaagcacttatttttcaatgctcagtgccagcgtg 5045

Search completed: April 22, 2002, 22:50:48  
 Job time: 27556 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:43:41 ; Search time 293.81 Seconds  
(without alignments)  
2791.181 Million cell updates/sec

Title: US-09-873-409-14

Perfect score: 3621

Sequence: 1 ttccgcttcgtcgtgact.....atgcacagtcagtcagtgta 3621

Scoring table: IDENTITY-NUC  
Gap 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/lna/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/lna/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/lna/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/lna/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/lna/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1316.2	36.3	4646	1	Sequence 2, Appl
2	1316.2	36.3	6505	2	Sequence 5, Appl
3	1316.2	36.3	9318	2	Sequence 6, Appl
4	1314.6	36.3	4669	6	Patent No. 5206352
5	1308.2	36.1	4669	2	Sequence 18, Appl
6	1308.2	36.1	4669	2	Sequence 1, Appl
7	1305.2	36.0	4264	2	Sequence 1, Appl
8	1303.6	36.0	4264	2	Sequence 5, Appl
9	1212	33.5	4233	4	Sequence 1, Appl
10	1212	33.5	4233	4	Sequence 1, Appl
11	856	23.6	2726	1	Sequence 1, Appl
12	496.8	13.7	4002	2	Sequence 1, Appl
13	496.8	13.7	4002	2	Sequence 1, Appl
14	496.8	13.7	4002	4	Sequence 1, Appl
15	496.8	13.7	4002	4	Sequence 1, Appl
16	437.4	12.1	4047	2	Sequence 3, Appl
17	392.2	10.8	4224	2	Sequence 1, Appl
18	365.4	10.1	4800	2	Sequence 1, Appl
19	352.6	9.7	3924	2	Sequence 3, Appl
20	195.6	5.4	3927	1	Sequence 1, Appl
21	195.6	5.4	3927	2	Sequence 1, Appl
22	194.8	5.4	3924	2	Sequence 1, Appl
23	194.8	5.4	3924	3	Sequence 3, Appl
24	181.6	5.0	3909	1	Sequence 3, Appl
25	149	4.1	2061	4	Sequence 1, Appl
26	145.2	4.0	2407	4	Sequence 17, Appl
27	145.2	4.0	2407	3	Sequence 2, Appl

## ALIGNMENTS

28	145	4.0	5120	3	US-08-772-270A-6	Sequence 6, Appl
29	145	4.0	8370	2	US-08-488-706-1	Sequence 1, Appl
30	144.8	4.0	7721	3	US-08-772-270A-14	Sequence 14, Appl
31	141.2	3.9	1959	4	US-09-061-764A-4	Sequence 4, Appl
32	133.4	3.7	1201	3	US-08-961-083-189	Sequence 189, App
33	133	3.7	2244	4	US-09-061-764A-18	Sequence 18, Appl
34	130.4	3.6	4403765	4	US-09-103-840A-2	Sequence 2, Appl
35	128.6	3.6	2376	1	US-08-394-880B-1	Sequence 1, Appl
36	128	3.5	4403765	4	US-09-103-840A-2	Sequence 2, Appl
37	128	3.5	4411529	4	US-09-103-840A-1	Sequence 1, Appl
38	117.6	3.2	6143	1	US-08-612-521-3	Sequence 3, Appl
39	110	3.0	28804	2	US-08-592-874-1	Sequence 1, Appl
40	110	3.0	28804	3	US-09-096-942-2	Sequence 2, Appl
41	110	3.0	28804	3	US-09-096-867-2	Sequence 2, Appl
42	104.2	2.9	5889	1	US-08-463-092B-5	Sequence 5, Appl
43	104.2	2.9	5889	2	US-08-462-109A-5	Sequence 5, Appl
44	104.2	2.9	5889	2	US-08-460-907B-5	Sequence 5, Appl
45	104.2	2.9	5889	3	US-08-463-179A-5	Sequence 5, Appl

RESULT 1  
US-08-181-471-2  
Sequence 2, Application US/08181471  
Patent No. 5641508  
GENERAL INFORMATION:  
APPLICANT: Li, Lingna  
APPLICANT: Lishko, Valery K.  
TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Thomas Fitting  
STREET: 12526 High Bluff Drive, Suite 300  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92130  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08181,471  
FILING DATE: 13-JAN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
REFERENCE/DOCKET NUMBER: ANT0029P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-792-3680  
TELEFAX: 619-792-8477  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4267  
US-08-181-471-2

Query Match 36.3%; Score 1316.2; DB 1; Length 4646;  
Best Local Similarity 62.3%; Pred. No. 0;  
Matches 2178; Conservative 0; Mismatches 1268; Indels 51; Gaps 5;

QY 162 caggttaccctgtatctatgttgaaatagtgctgcttgatcttcttgatcatca 221  
DB 760 CAGGTATCCCATATATATACAGTGAATGCTGCTGGGTGCTGGTGGCTGCTTACATCA 819

QY 222 gatttcctgttgatataactgacagcagcagcagcagcagcagcagcagcagc 281  
DB 820 GGTTCATTTTGGTGGCTGGCAGCTGGAAGCAATACAAATTTAGAAACAGTTT 879

QY 282 tcaatcagtttgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 341  
DB 880 TCATCTATATATGCGACAGAGATAGGCTGGTTGATGTCACGATGTTGGGAGACTTAA 939

QY 342 cactgcagatgacagacatcagc---aaatcagtgatggtatgagatgaattgctc 398  
DB 940 CACCGACTTACAGATGATGCTCTTAAGATTAAATGAGTTATGGTGACAAATTTGGAAT 999

QY 399 gttgttcaaaaacatgctacttcttgatggcctgagcagtggttggtgaagcgtg 458  
DB 1000 GTTCTTTAGTCATATGCAACATTTTTCAGTGGTTTATAGAGATTATACAGTGTTG 1059

QY 459 gaactccaccctagtgactctatccacgctcctctataatgagcagcagcagcagc 518  
DB 1060 GAGCTAACCCGTGTGATTTTGGCCATCAGTCTGTCTTGAGACGTGAGCTGCTG 1119

QY 519 ttctagatgylcalctcaatgacagtaagaatlaagtgccattccaaaagctgggc 578  
DB 1120 GGCMAAGTACTATCTTATTTACTGATTAAGAACTTAAAGCTTAAAGCTGAGAC 1179

QY 579 tgtgcaagaagaagctgtgatacaatccgaacagcagcagcagcagcagcagcagc 638  
DB 1180 AGTACTTAAGAGGTCTTGCGCAGCATTAAGACTGTGATGATTTGAGCAAAAGAA 1239

QY 639 agaactcaaaagatatacagaatctcaagaatcaagaatcttgatcttgatcaaaaagg 698  
DB 1240 AGAAGCTGAAGGTACACAAATAATTAGAAAGCTAAAGAAATTTGGATTAAGAAAGC 1299

QY 699 tatagctcaaaagtgctctgtgctgctgctgctgctgctgctgctgctgctgctg 758  
DB 1300 TATTACACGCAATATTTTATAGGTGCTGCTTCTGATCTATGATCTTATGCT 1359

QY 759 tgccttggatggaacccctgattcttcaatgagaacccctgatatccatccagcagc 818  
DB 1360 GGCCTTCTGATGAGGACCACTTGTCTCTCAGGGG---ATATTTCTATTGGACA 1413

QY 819 tcttctgctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 878  
DB 1414 AGTACTCTACTATCTTCTTCTGATTAATTGGGGCTTTAGTGTGGAAGGCAATCTCC 1473

QY 879 tcaacttgaacccctgcaatagcccgagagcgtgcttccatcttccaggttatgga 938  
DB 1474 AAGCATTTGAACCATTTGCAAAATGCAAGAGGACCTTATGAAATCTTCAAGATATTGA 1533

QY 939 taagaaccccgatagataacttccacagcagcagcagcagcagcagcagcagcagc 998  
DB 1534 TAAATAGCAAGTATGACAGCTATTTGCAAGAGTGGCAAAACCAAGATTAATTAAGGG 1593

QY 999 aacttggaaatttaaaaagtcttcttcaatlatccatcaagaacatcatcaatctct 1058  
DB 1594 AAATTTGAATTCAGAAATGTTCACTTCACTTACCATCTGGAAGAAAGATTAATCTT 1653

QY 1059 gaaaggttcagatctcagaatlaagctcgtgagagacagcagcagcagcagcagcagc 1118  
DB 1654 GAAGGCGCTGGAACCTGGAAGGTGCAAGGTGGCGAGCGCTGCTGTTGGAAGACAGTGG 1713

QY 1119 cagtgaggaagagtagcagtagcagcagcagcagcagcagcagcagcagcagcagc 1178  
DB 1714 CTGTGGGAAGAGCAACACAGTCCAGCTGATGACAGGCTCTATGACCCACAGAGGGGAT 1773

QY 1179 tatcatggtgagatgagaatgacatcagagcctttaaagtgcgcatctcagacacat 1238  
DB 1774 GGTCAAGTGTATGACAGATATTAAGACATTAATGAAGTTTATACAGGAATCAT 1833

QY 1239 tggagtggttagtcaagagcctgttctgttcggagcccaatcagtaacatlaagta 1298  
DB 1834 TGGTGTGTGATGATGAGCACTGATTTGTTGCCACCAAGATAGCTGAACCATTCGCTA 1893

QY 1299 tggacagatgatatgcatgatatgaagatgagaagcagaagaggaagcagaatgcta 1358  
DB 1894 TGGCGGTGAATATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGAT 1953

QY 1359 tgatttatacagaggttctcctaataaatttaatacatggtgaggggaaaagagctca 1418  
DB 1954 TGACTTATATGAAAGTCCCTCATTAATTTGACACCTGTTGGAAGAGAGGGGCCCA 2013

QY 1419 aatgagtgagggagcagaacagagatgagcagcagcagcagcagcagcagcagcagc 1478  
DB 2014 GTTGAAGTGGCGAGAGAGAGATGCGCATTTGACCTGCTGCTGCTGCTGCTGCTGCTG 2073

QY 1479 gattcagattttagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1538  
DB 2074 GATCTCTCTGCTGAGTGAAGCCACGCTCAGCTTGGACACAGAAAGCAGAGCTGTCA 2133

QY 1539 agctgcactggaagaagcgaagcagaagcagcagcagcagcagcagcagcagcagcagc 1598  
DB 2134 GGTGCTCTGATAGAGCGACAGAAAGGTGCGACCACTTGTGATGCTCATCGTTGTC 2193

QY 1599 tactcttgaagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1658  
DB 2194 TACAGTGTGATAGTGTGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2253

QY 1659 agcacatctgaactaagcagaagcagcagcagcagcagcagcagcagcagcagcagcagc 1718  
DB 2254 AAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2313

QY 1719 tattaanaaagctgatacagcagatgagatgagatgagatgagatgagatgagatgagat 1762  
DB 2314 AGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2373

QY 1763 -ctgaagaagaagcagaactcaactctcctgcatctgagagagacata----- 1810  
DB 2374 CTGGAAATGCTCTTCAATATGATCAAGATCCAGCTTAATTAAGAAAGATCAACTGTAG 2433

QY 1811 -----agtcaacttcatgacaagcgtgagaatccacacacatgaagat 1859  
DB 2434 GAGTCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2493

QY 1860 aagcttccctgaagctctctataaaaaatttaaaaglttaaaacaagcctgaatgctct 1919  
DB 2494 AAGTATACCTCCAGCTTCTTGGAGGATTAAGCACTTAATTAATTAATTAATTAATTAATTAAT 2553

QY 1920 tgtggttctgggagatgagcctgctgcttcaatgagaactgltatccagatattccat 1979  
DB 2554 TTTTGTGTTGGTGTATTTTGGCCATTAATTAAGAGGCTGCAACCAAGCATTTGCAAT 2613

QY 1980 catcttgcagaatataaacaatgltt---ggaataatgataaaacacatlaaaga 2036  
DB 2614 AATATTTTCAAAAGATTAATAGGGTTTTCAGAAATTAATTAATTAATTAATTAATTAATTAAT 2673

QY 2037 tgalgcagaatltatccatgatatctgcatatttggtggttatttgcttgatgta 2096  
DB 2674 GAATGTAACTTTGTTACATATTTGTTCTAGCCCTTGGAATTAATTTATTAATTAATTT 2733

QY 2097 ttcatgcagagatatttccgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2156  
DB 2734 TTTCTTCAAGGTTTCACTTTTGGCAAGCTGGAAGATCTCCACCAAGCGCTCCGATA 2793

QY 2157 ctggccttcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2216  
DB 2794 CATGTTTTCGATCAGTCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2853

QY 2217 cacaagagcctgacaacataatagcagatagatagacaaatccaagagcaagcagc 2276

Db 2854 CACTGAGACATTTGACCTACAGGCTCGCCATATGCTGTCTCAAGTTAAAGGGGCTATAG 2913  
Qy 2277 ttccaggaattgagcgtcttaacacaaatgcaactaacatgaggaacttcaagtatcattc 2336  
Db 2914 TTCCAGGGCTGTCTGAATATACCCAGAAATATAGCAATCTTGGACAGGAATATATATTC 2973  
Qy 2337 cttatataatgaggaatgagcaatctcgtatcttgatgttcgaactcgaacttgcgct 2396  
Db 2974 CTTCAATCTATGCTTGGCACTACACTGTTACTCTTAGCAATTTGACCATATTTGCAAT 3033  
Qy 2397 gaccaggaatgataaaccgcagcaatgactgattctgccaacaaagaataaagaact 2456  
Db 3034 AGCAGAGCTGTTGAAATGAATAATGTTCTGACAGACACAGTAAGTAAGTAAGTAAGTAAG 3093  
Qy 2457 taagcaatgctggaagaataagcaactgaagccttgagaataatagctactatgctcatt 2516  
Db 3094 AGAAGGCTGCTGGGAAGATCGCTACTGAAAGCAATGAAATCCGACCGTTGTTCTTT 3153  
Qy 2517 aacaagggaagaagccttcgagcaatgataagaagatgcttcgaactcaacagaa 2576  
Db 3154 GACTAGAGGAGCAAGATTGAAATATATATGCTCAGAGTTGACAGTACATACAGAA 3213  
Qy 2577 taacctgaagaagaacagatatttggaagcctgattgacattcaagcattatata 2636  
Db 3214 CTCTTTGAGGAAGACACACATCTTGGAAATTTACATTTCTTCCACCAAGCAATGATGTA 3273  
Qy 2637 ttctgactatgacagaggttgcatttggaacttaattcaagcctgagcaagcagc 2696  
Db 3274 TTTTTCATATGCTGATGTTTCCGTTTGGAGCCCTACTGCTGAGCAATTAATCATCATGAG 3333  
Qy 2697 cccaaaggagcgttcattcatttctacgcaattgataagatgagcattgagcattgagaa 2756  
Db 3334 CTTTGAAGGATGTTCTGTTAGATTTTACGCTGTTCTTTGGTGCATGCGCGTGGGCA 3393  
Qy 2757 aacgctcttctgctgcaatlatccaaacaaatcgcgaggtgctgcaacttctgc 2816  
Db 3394 AGTCAGTTCATTTGCTCCGACTATGCCAAGCAAAATATACAGACCCCATCATCATCAT 3453  
Qy 2817 ctgtctggaagaagaacaaatataagacagccgagctcaagaagggaagaacagacac 2876  
Db 3454 GATCATTTGAAAGAACCCCTTGTGATGACAGCTACAGCAAGGAGCTTAATCCGCAACAC 3513  
Qy 2877 atgtgaagggaattgaagtttcgagaagctcttcttctcattcattgcgcagatgt 2936  
Db 3514 ATTGGAAGGAAATGTCACATTTGATGAGTTGATTCATTCATCCACCCGACGACAT 3573  
Qy 2937 ttctcatcccgctgctatccctcagtaattgacgaggaagaagaagatgattgtg 2996  
Db 3574 CCCACTGCTTCAGGACCTGAGGCTGAGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 3633  
Qy 2997 gaggcagcgctgctggaagaagaacacttctgcaactctgcagagacttataagccgct 3056  
Db 3634 CAGCAGTGGCTGTGGAGAGACAGAGTGTCCAGCTCCGAGGCGGTTCTAGACCCCTT 3693  
Qy 3057 gcaaggacaaggtcgttctggtgctgctgcaagaagaagaatggaatgacagtgctgcg 3116  
Db 3694 GGCAGGGAAGGCTGCTGATGAGGCAAGAAATGAAGGAGCAATGTTCACTGCTGCTCG 3753  
Qy 3117 ttcccaaatagaatcgttccctcagaagcctgctcttcaactcagactgctgaga 3176  
Db 3754 AGCACACTGGGCACTGCTGCTCCAGAGGCCATCTGTTGACTGACATGCTGAGAGA 3813  
Qy 3177 catgcctatagtgacaacagcgtgctgctcattagatagatgataaagaacgcgcgca 3236  
Db 3814 CATTCCTATGAGAGACAGACCGGGGTGTACAGAGAAAGATCGTAGGCAAGCA 3873  
Qy 3237 tgcagaataatcacttcttcttcttctgaggtctcctgagaataacacacagttg 3296  
Db 3874 GGAGGCCAACATACATGCTTATCATGAGTCACTGCTTATTAATATAGCACTAAAGTAG 3933  
Qy 3297 actgaaagagacagccttctgctgagcagaaacaaagcactgacttgcagagcct 3356

Db 3934 AGACAAGAACTCAAGCTCTGTGTGGCCAGAAACACGCAATTTGCCATAGTCTGTCCCT 3993  
Qy 3357 tctccaaaaaaccaaaatttatttgttgatgagggcaacttaagccctcagataatgacag 3416  
Db 3994 TGTTAGACAGCTCTATTTTCTTTTGGATGAACCCAGCTAGCTGCTGGAATACGAAG 4053  
Qy 3417 tgaagagtgctgctgagcctccttgataaagcaggaaggaagatgctgctagtgt 3476  
Db 4054 TGAAGAGTTGTTCAGAGAGCCCTGACAAAGCCAGAGAGGCGCACCTGCTGCTAT 4113  
Qy 3477 cactcagagcctctgcaatcagaacgcaatgataagtgctgctgcaacatgaa 3536  
Db 4114 TGCCTACCGCTGTCCACCATCAGAAATGCACTTAATATGATGTTTCAAGATGGAG 4173  
Qy 3537 gataaaggaacaaactcaagagctcctgagaatcgagacataatttaagt 3596  
Db 4174 ACTCAAGGAGCTGACACCATCAGCTGCTGACAGAAAGCAATCTATTTTCAAT 4233  
Qy 3597 agtgaatgacagcag 3613  
Db 4234 GGTCACTGTCCAGGCTG 4250

RESULT  
2  
US-08-793-610-5  
; Sequence 5, Application US/08793610  
; Patent No. 5858744  
; GENERAL INFORMATION:  
; APPLICANT: BAUM, Christopher  
; APPLICANT: STOCKING-HABERS, Carol  
; APPLICANT: OSTERAG, Wolfram  
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
; TITLE OF INVENTION: FOR GENE TRANSFER  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikolaide, Marmelstein, Murray & Oram LLP  
; STREET: 655 Fifteenth Street N.W. Suite 330  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/793,610  
; FILING DATE: 07-MAR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 31 973.8  
; FILING DATE: 08-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 195 03 952.1  
; FILING DATE: 07-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP95/03175  
; FILING DATE: 10-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Berman, Richard J.  
; REGISTRATION NUMBER: 39,105  
; REFERENCE/DOCKET NUMBER: P1614-7007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)638-5000  
; TELEFAX: (202)638-4810  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6505 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA  
; US-08-793-610-5





QY	2217	caagagggcgttacacaacatcttagccaatagatlagacaacaaatccaagagcaagc	2276
Db	4246	CACGTGGACGCTTGAACACACAGGCTCCGCAATGATGCTGCTCAAGTTAAAGGGCTATAGG	4305
QY	2277	ttccagagatttggcgctcttaacacaaaatgcaactaacatcagagacttcagttatcatctc	2336
Db	4306	TTCCAGGCTTGCGTGAATATACCCAGAAATATAGCAAACTGTGGACAGGAATATATATAC	4365
QY	2337	cttatatatgabatvggagatgacatctccatgattctgtagtattgtctcagtaactgtccgt	2396
Db	4366	CTTCTATGTATGGTTGGCACTACACTGTTACTCTTACGAATATGTATACCATCATGTCAAT	4425
QY	2397	gaacagatgatattgaaacccgacgatgtcctgtgattttgcacaagaatagaagaaact	2456
Db	4426	AGCGAGGTTGTTGAAATGAAATGTTGTGTGGACAGGACTGAAATATAAAGAACT	4485
QY	2457	taagcagctcgtgaagaatagacaactgaaactcttgagaaatagactactatagttaact	2516
Db	4486	AGAAAGTCTGGGAAAGATCGCTACTGACCAATAGAAAACTTCCAAACCGTTGTTCTTT	4545
QY	2517	aacaagggaaaaagcccttcgagcaaatglatgaagagatgctcagactcaacacagaa	2576
Db	4546	GACTCAGGACGACAAAGTTGAACATATGATGTCAGATTTGCGAGTATACATACAGAA	4605
QY	2577	taactcgaagaaagcacagattatctggaagcgttatatgcatcagccatgtccttata	2636
Db	4606	CTCTTTGAGGAAAGCACACTTCTTGGAATATACATTTTCTTACCCAGGCAATGATGTA	4665
QY	2637	tttgcctcagcagcaggggttcgatcttgagaccttaattcaagctgagcgaataac	2696
Db	4666	TTTTTCTTATGCTGGATGTTTCCGGTTTGGAGCCTTACTTGGTGCAATAACTCATAG	4725
QY	2697	ccccaggggacatgtlcatagtttcttaactgcaattgagatagagcatalgccccagaa	2756
Db	4726	CTTTGAGAGATGTTCTGTAGTATTTTCACTGTTGTCTTGGTGCAATGGCCGTGGCA	4785
QY	2757	aagcgtcgttttgctctccgtaatatccaaagcaaaatcgggggctgcgacatctgttc	2816
Db	4786	AGTCAGTTCACTTGTGCTCTGCACTATCCAAAGCAAAATATAGCAGCCCATCATCAT	4845
QY	2817	cttgttgaaagaagaaccaaatctagaagcccgacgtcaagaagggaagaagccagaac	2876
Db	4846	GATCATTTGAAAAAACCCTTTGTTATACAGCTACACACGAGAGGCTTATGCGCAACAC	4905
QY	2877	atgtgaaggaattcagaagtttcgaaagctctctctctctctatccatgtgcgcagact	2936
Db	4906	ATTGGAAGGAAATGTCACTATTGTGTGAAGTGTATTCACTATCCACCCAGCCGAGAT	4965
QY	2937	tttcatctccgtgtgcttaccctcagatctgagcagagaaagaagatgacatctgtg	2996
Db	4966	CCCAGTCTCTCAGGGACTGAGCTGGAGGTGAAGAAAGGCCAGACGCTGTGTTGGG	5025
QY	2997	gagcagcgctgtgtgggaaagaacactctgttcaactcttcgagagacttatgacccgt	3056
Db	5026	CAGCAGTGGGTGTGGGAAGGACACAGTGTGCCAGCTCTCGAGGCGGTTCTTAGACCCCTT	5085
QY	3057	gcaagagcaagtgctgttttatgtgtgtagtgaagaagattgaaatgtaacatgtgctcg	3116
Db	5086	GGCAGGGAAAGTCTGCTTGTATGGCAAAAGAAATAAAGCACATGANTGTTCAGTGGCTCG	5145
QY	3117	tttcccaaatgcaatcgttcttccaaagccctgtctctcaactcgaagcatgtcgtgaa	3176
Db	5146	AGCACACCTTGCGATCGTGTCCAGAGACCCATCTGTTGTACTGCACGATGTGTGA	5205
QY	3177	catcgctatgtgtgaacaagccgtgtgtgtgcattagtgtgtagtcaagaagcccgaa	3236
Db	5206	CATTGCTTATGGAGACAAACGCCGGGTGTGTACAGGAAGATCTGTAGGGCAGCAAA	5265
QY	3237	tgcagcaaatatcatctcttcttcttgaagtgctcccttggaatatcaacacacaagtttg	3296
Db	5266	GGAGGCCAACATACATCCCTTCAATCGAGTCACTGCTATTAATATATAGCACTAAAGTAGG	5325
QY	3297	actgaagagacacgcttcttgctgctgccaagaacaaagactagcatatgcaagggctct	3356

Db	5326	AGNAAAGGAAACCTACGCTCTCTGTGTGGCCAGAAACAAGCCATTGGCCATAGCTCGGCCCT	5385
QY	3357	lclccaaaaaaccaaaatlattlgtltgagtgagggccactlccagccctcgataalagacag	3416
Db	5386	TGTTAGACACCTCATATTGCTTTGGATGAAGCCACGTCAGCTCGATACAGAAAG	5445
QY	3417	tggagaagtgttccagatgcagtccttataagaagcgaagggaaagaaatgcctagtgt	3476
Db	5446	TGAAAAGTTGTCCAAAGAAACCCCTGGACAAAGCCACAGAAAGCCGCCACCTGCATTGTAT	5505
QY	3477	cactcacagagctctctgcaatlccagaacgaaatltagatagtgtctgcacaatlgaagaa	3536
Db	5506	TGCTCACCGGCTGTCCACACTCCAGATCCAGACTTAATAGTGGTTCAGAAATGCAG	5565
QY	3537	gataaaggaacaaagaaactlcaatlaagagctcctgagaatcgaagacatatalattlaagt	3586
Db	5566	AGTCAAAGGACATGGCCACGATCAGCATGCTGTGGACAGAAAGGACATCTATTTTCAAT	5625
QY	3597	agtgaatgcacagtlcaag	3613
Db	5626	GGTCAGTGTCCAGGCTG	5642

RESULT 3  
 US-08-793-610-6  
 Sequence 6, Application US/08793610  
 Patent No. 5858744  
 GENERAL INFORMATION:  
 APPLICANT: BAUM, Christopher  
 APPLICANT: STOCKING-HARBERS, Carol  
 APPLICANT: OSTERTAG, Wolfram  
 TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
 TITLE OF INVENTION: FOR GENE TRANSFER  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nikolaio, Marmelstein, Murray & Oram LLP  
 STREET: 655 Fifteenth Street N.W. Suite 330  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005-5701  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/793,610  
 FILING DATE: 07-MAR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE P 44 31 973.8  
 FILING DATE: 08-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: DE 195 03 952.1  
 FILING DATE: 07-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP95/03175  
 FILING DATE: 10-AUG-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Berman, Richard J.  
 REGISTRATION NUMBER: 39,105  
 REFERENCE/DOCKET NUMBER: P1614-7007  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)638-5000  
 TELEFAX: (202)638-4810  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9318 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA

US-08-793-610-6

Query Match	36.3%;	Score 1316.2;	DB 2;	Length 9318;
Best Local Similarity	62.3%;	Pred. No. 0;		
Matches 2178;	Conservative	0;	Mismatches 1268;	Indels 51; Gaps 5

OY	162	caagcttaccctcctatcattcgttggaataagctcttgcctccgtgatttttgatataca	221
Db	2111	CAGGTAATGGCTATATATACAGTGGAAATTCGTCTCTGGGGTGGCTTGCCTACATTC	2170
OY	222	gatttccttgtagtataatactgcagcacagacacaaaggatctcgaaaacatttt	281
Db	2171	GGTTCATTTTGGTGGCTGGCAGCTGGAGCTGGAAACAAATATACAAATATAGAAACAGTTT	2230
OY	282	tcattcagtttggccacgaacatcgcgctggttttgaatagctctgacatcggttaactta	341
Db	2231	TCATGCTATAATATGCGACAGAGATATAGCGTGGTTGATGTGCACAGATTTGGGAGCTTA	2290
OY	342	cactgcacatgcagacatcagac---aaatcagltgatgtgatgaataaattgctc	398
Db	2291	CACCGACACTTACAGATGATGATGCTCTAAGATTATATGAAGTATTGTGGACAAATTTGGA	2350
OY	399	gttgttccaaacatgctacttcttcagatltggccctgcagcttggttttgtaaggctg	458
Db	2351	GTTCCTTACGTCAATGGCAACATTTTTCACCTGGGTTATATGTAGGATTTACACGTGGTTG	2410
OY	459	gaacttcacccctgtaactatccacagctcctcctctataatgagcttcacgcagcagatg	518
Db	2411	GAAGCTAACCCCTTGATGTATTTGGCCATCACTCCTGTTCTTGACACTGTCACTGCTGTCTG	2470
OY	519	cttcagatgtgcatcctcatctgacccagtaaggaaatlaagtgccbatlccaaagctgggc	578
Db	2471	GGCAAGATACATATCTTCATTTACTGATTAAGAACTCTTACGATATGCAAAAGCTGAGC	2530
OY	579	tgtgacgaagaagctctgtgcatacaatccgaagaagatagatccctttagggcccgagaa	638
Db	2531	AGTAGCTGAAGAGCTCTTGCGACGAATTTACATGCTGATTTGACATTTGGAGACAAAAGAA	2590
OY	639	agaacttcaaaagctatacacagaaatctcaaaagatgcacaaagatttggcataaaaggac	698
Db	2591	AGAACTTGAAAGGTACAAACAAATTTAGAAAGATCTAAAGAAATTTGGATTAAGAAAGC	2650
OY	699	tatagcttcaaaagctcctcttggtgtctgtgtacttcttatgaaatggaaactatlgact	758
Db	2651	TATTACAGCCAAATATTTCTATAGTGTGCTGCTTCCTCTGATCTATATCATCTTATATGCTCT	2710
OY	759	tgccttttgatagaaccctcccttgatctcttaatggaagaaacttgatataccacggagac	818
Db	2711	GGCCTTGTGGTATGAGGACCACTTGTCCTCTCAGGGGA-----ATATCTATATTGGACA	2764
OY	819	tgctctgcctgttctctttagtgaatccatagacagtatgtcatltggagcagcagctcc	878
Db	2765	AGTACCTACAGTATCTTTTCTGTATTAATTTGGGGCTTTTATGTGTGACAGCATCTCC	2824
OY	879	tcactttgaaaccttgcacaatagcccgagagagcttgcttcttataatlttccaggtatgta	938
Db	2825	AAGCATTTGAAGCATTTGCCAAATTCGACAGGAGCAAGCTTATGAATCTTCAAGATPATTTGA	2884
OY	939	taagaaacccagatagataaacttcttccaaagccggagataaacttgatcatagaaag	998
Db	2885	TAAATAGCCAAAGTATTGACAGCTTATTCGAAGATGGGACAAACCAAGATATATTAAAGG	2944
OY	999	aacttggaaatttcaaaatgttcttcttcaattatccatcaagaacatctatcaaatctc	1058
Db	2945	AAATTTGGAATTCAGAAATGTTCACCTTCAGTTACCCTCGAAAAAGAAAGTTAAGATCTT	3004
OY	1059	gaagatgttgatctcgaatlaagtctcgaggagaaacagctgccttggtgcgtgtctcaatcgg	1118
Db	3005	GAAGGGCTTGAACTGGAAGGTGCAGAGTGGCGACAGCGGTGGCCCTGGTTGGAAACAGTGG	3066
OY	1119	cagtgaggaaagtaagctagctcagagctctcgaagaggttatatgatlccggatgacgctc	1176

Db	3065	CTGGGGAAGACACAACACTCCAGCTGATGCAGAGGCTCTATTGACCCCAAGAGGGGAT	3124
Qy	1179	tatacgtgtagatgagatgacatcgagcctttaaattgtgcgcatlatacgagccat	1238
Db	3125	GGTCAGGTTGATGGACAGGATATTAGGACCAATTAATGTAAAGTTCTTACGGGAATCAT	3184
Qy	1239	tggagtgttagtcaagaagcctgtttgttcggagccaccatcagtaacaatatacagta	1298
Db	3185	TGGGTGGTGAGTCAGCAACCTGTATTGTTTGGCACCAAGATAGCTGAACCAATTCCGCTA	3244
Qy	1299	tggagagatgatactgacatgaaagatgtagagagcagcaagaggaagcaaatlycgt	1358
Db	3245	TGGCCGTAAAAATCTCCACATGATGATGATTTGGAAGAAGCTGTCAAGGAAGCAATGCTTA	3304
Qy	1359	tgaattlatactgtaggttccctaataattlaatacatgtgtgaagggaagaagctca	1418
Db	3305	TGACCTTATCATGAAACCTGCCTCATTAATTGTACACCCCTGGTTGAGAGAGGGGCCCA	3364
Qy	1419	aatgtaggagggcagaacaagagatcgcaattgtctgtgcttaagttcgaaccocaa	1478
Db	3365	CTTTCAGTGGTGGCAGAAAGCAGAGGATCGCATTGGACGTCCGTGGTTCGCAACCCAA	3424
Qy	1479	gattctgtatttagatgagtgatgctgtgccggattcgaagaagaatgcagctgttca	1538
Db	3425	GATTCCTCTGCTGGATGAGGCGCACAGTGCCTTGGACACAGAAAGGAGCAGTGTTTCA	3484
Qy	1539	agctgcatctgtagagagcgagcaaaagtcgagatataatgctgtgttagcacaccgaattc	1598
Db	3485	GGTGGCTCTGATTAAGCGCCAAAAGGTGGAGCACACCATTTGTATGCTCATTCGTTTTC	3544
Qy	1599	tactatcgaaagtcagatcttgatttgcacctaaagatggagatgtcgtggcggaagaag	1658
Db	3545	TACAGTTCGTATGCTGACGTACGTACGCTGGTTTCGATGATGGAGTCAATTGTGAGAAAG	3604
Qy	1659	agcaatctctgacatctgagcaaaagcgctatatacttaccctgtgattgcagga	1718
Db	3605	AAATCTGTGATCACTCATGAAAGAAAGGATTTACTTAAACTTGTCCATCATGAGAC	3664
Qy	1719	tattaaaaaagctgtagacagatgtagtaatgacatltcta-----	1766
Db	3665	AGCAGAAATGAAGTTGAATTAGAAAAAGCGAGTGAATCCAAAAGTAAATTGATGC	3724
Qy	1763	-ctgaagaagaagcaactactctctctgacatctgtgaaaggatca-----	1810
Db	3725	CTTGGAAATGCTTCAAAATGATTCAGATCCAGTCTAATAAGAAAGATCAACTCTGAG	3784
Qy	1811	-----agtcagactcatgacaaagcctgaggaatccacccaactaaagat	1855
Db	3785	GACTGTCCGTGATCACAGCCCAAGACAGAAAGCTTAATGCCAAGAGGCTCTGATGA	3844
Qy	1860	aagctctctgtagtctctctataaaaaattlaagltaaacaagcctgaatgctctt	1911
Db	3845	AAATTAACCTCCAGTTTCTTTTGAGAGATTATGACGTAAATTTAACTGAATGGCCTTA	3900
Qy	1920	tgtgtgtctgggacatgtggtctctgtcttaatggaacgtgtcatcagtatlttccat	1971
Db	3905	TTTTGTTGTGGATTTATTTTGTGCATTTAAAGTGAGGCTCGAACACACATTTGCAT	3966
Qy	1980	caacttgcataaataatatacagtgttt-----ggaataatgatanaacccactaaaga	2033
Db	3965	AAATATTTTCAAAACATTAATATGAGGGTTTTTTCACAAAGATGATGATCTCTTAACAAACAGACA	4022
Qy	2037	tgtatgcagaatattatccatgatactgcgtcatatttgglytlatattgtcttgagta	2099
Db	4025	GAATAGTAACCTTCTTTTCACTATTGTGTCTAGCCCTTGATATTTCTTTTATTCATTT	4088
Qy	2097	ttcatatgcaggatatttllacgcaagagcaagggaatttlaacgtatgatttaagac	2155
Db	4085	TTTCCCTTACAGGTTTACATTTTGGCAAAAGCTGGAAGATCTCCACCAAGGGCTCCGAT	4144
Qy	2157	cttggccttaagagcatgtlatacagagatactgcgglytlatgtaaaaggaacag	2211
Db	4145	CATGCTTTCCGATCATGCTCAGACAGAGATGTAGTTGGTTTATGATCAACCTTAACAAACAC	4200

Qy	2217	caacgagagccttgacaacaatalatgacatagatagacacaatatccaagaagcaacag	2216
Db	4205	CACtgcGAGcATTgACTTACACAGGCTCCCAATATATCTCGTCAAGTCTTAAAGGGCGTATAGG	4264
Qy	2277	tccaagattgagcgtcttaacacaatatgcaactacaataggaacttcagatcatctc	2336
Db	4265	TTCCAGGCTTGCTGTATTATACCGAANATATACCAATCTGGGACAGGAATATATTATATC	4324
Qy	2337	cttataatgatgagggagatgacatccctgatactgtagatctgcacgaactgcgcgt	2396
Db	4325	CTTTCATGTATGTTGGCACTAACACTGTATTACTTTAGCAATATGTACCATCATTTGCAAT	4384
Qy	2397	gacagcaatgatgtgaacccgacagcaatgacatgcttgatttcgcaacaaagataagcaagact	2456
Db	4385	AGCAGGAGCTTGTGAAATGAAATGTGTGCTGGACACAGCTGAAAGATTAAGAAAGACT	4444
Qy	2457	taagcaltgctgaaagataagcaactgaaagctcttggaagataacgtaactatagtgcat	2516
Db	4445	AGAAGTGCTGGGAGAGATGCCCTACTGAACCAATAGAAAACCTCCGAACGTTGTTCTTT	4504
Qy	2517	aacaaaggaaaaagccctcctgagcaaatgfatagaagaatagtcctgaactcaaacagaa	2576
Db	4505	GACTCAGGACACGAAGTTGTAAATATGTATGCTCAGATTGTGCAGGTACATACAGAA	4564
Qy	2577	taccctgaagaagacaagataatgtgaagcgttatgatcatcaagcaltgctctatata	2636
Db	4565	CTCTTTGAGGAAAGCACACACTCTTTGGATTTACATTTTCCTTACCCAGGCAATGATGTA	4624
Qy	2637	tttgacctatgacagcaagggttcgattctggagacctaatltaactcaagctlgacgaatgac	2696
Db	4625	TTTTTCTTATGCTGGAGATGTTCCGGTTTGGAGCTACTTGGGCACTAATACCTCATAG	4684
Qy	2697	cccaaggagcaatgtcaatgttttctacacatgtgataatlgagcctatgagccatcgaa	2756
Db	4685	CTTTGAGGATGTTCTGTAGTATTTTTCACGCTGTTGCTTGGTGCATGCGCGTGGCA	4744
Qy	2757	aacgctgcttctgctctgcaatattccaaagccaatcggggcgctgacatctgttgc	2816
Db	4745	AGTCAGTTCAATTGCTCTCCTACTATCCCAAGCAAAATATCAGACGCCACATCATCAT	4804
Qy	2817	ctgtgtgaaagaagaaccaaatatagaacgcccgatgcaagaaggaagaaagccaagaac	2876
Db	4805	GATCATTTGAAAAAACCCTTTGTATTAACGCTACACACGAGAAAGCCTTAATGCCAATAC	4864
Qy	2877	atgcaagaggaatliagaagttctcgagaagctctctctctctatccatgtgcgccagatgt	2936
Db	4865	ATTGGAAGGAATATGTCACTATTGTGGTAAGTGTATTCACATATCCACCCGAGCAAT	4924
Qy	2937	tttaactcctcgatggtcttataccctcaagtalatgacgaggaagaagacagtagcaattgtgg	2996
Db	4925	CCCAGTCTTCAAGGACTGTGACCTGTGAGGTGAAGAAAGGCCACGACGTCTGTGTGG	4984
Qy	2997	gagcagcggcgttgagaaagcaactctgttcaactctcgagagacttatgacccgt	3056
Db	4985	CAGAGTGGCTGTGGGAAGACACATGTGTCCAGCTCCTGGAGCGGTTCTACAGACCTT	5044
Qy	3057	gcaagaaacaaatgctgttgaatgagtgatgacaaagaatgaaatgtacagtggtctcg	3116
Db	5045	GGCAGGGAAGATGTCTGCTTATGTGCAAAAGAAATTAAGCGAGTGAATGTTCAGTGTCCG	5104
Qy	3117	ttccaataatgcaatgcttctcctaagaagctgagctctcaactcagcatgtgcgtgaa	3176
Db	5105	AGCAACACTGGCATCTGTGCCAGGAGGCCATCCTGTTTGACTGACGATGCTGTGA	5164
Qy	3177	catgcctatgttgaaacaacgccgtgtgtgcatcatagaatgactcaagaagcccgcaa	3236
Db	5165	CATTGCGTATGAGACAAACAGCCGGGGGTGTACAGGAAGAGATCGTAGGGCAGCAA	5224
Qy	3237	tgcaagcaaatatcatctcttatatgaaggtctcccttgagaaatacaacaacaagtgg	3296
Db	5225	GGAGGCCAACATATATCTTCAATCGAGTCACTACCTCTAATTAATATAGCACTAAATGTGG	5284

QY	3297	actgaagagacagcgccttcttgcggccagaaaagaactagctatctgaagggcctc	3356
Db	5285	AGACAAGGAATCTACGCTCTCTGGGGCCAGAAACAGCATTTGCCATATGCTCGGCCCT	5344
QY	3357	tctccaaaaccacaatttatctgtgtgataaggccacttcagcccccgaatgacag	3416
Db	5345	TGTTAGACAGCCTCATATTTTGGCTTTTGGATGAGACCCACGTACGCTCTGGATACGAAAG	5404
QY	3417	tgaagaagtggttcagcagctgtcgtataaagccagacagcggaaggaatgcctagtg	3476
Db	5405	TGAAAAGGTGTTCCAGAAAGCCCTGGACAAAGCCAGAGAGAGCCCGCACCTGATTTGTAT	5464
QY	3477	cactcagagctctctgcaatcagaacgcagatttgatgtgtctgcacaaatgaaa	3536
Db	5465	TGCTCACCGCCCTGTCACACATTCAGAAATGACACTTAATAGTGCTTTACGAAATGGCAG	5524
QY	3537	gataaaggaacaggaactcattcaagagctcctctgagaatcgaagcatalatttaagt	3596
Db	5525	AGTCAGAGGACATGGACGACATCACAGCTGCTGGACAGAAAGCATATTTTTCAT	5584
QY	3597	agtgatgcacagtcag 3613	
Db	5585	GGTCAGTGTCCAGGCTG 5601	
RESULT 4			
5206352-3			
Patent No. 5206352			
: Applicant: Robinson, Igor B.; Pastan Ira H.; Gottesman,			
: Michael M.			
: TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA			
: SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS			
: NUMBER OF SEQUENCES: 4			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/07/622,836			
: FILING DATE: 24-SEP-1990			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: 892,575			
: FILING DATE: 01-AUG-1986			
: APPLICATION NUMBER: 845,610			
: FILING DATE: 28-MAR-1986			
: SEQ ID NO:3			
: LENGTH: 4669			
5206352-3			
Query Match 36.3%; Score 1314.6; DB 6; Length 4669;			
Best Local Similarity 62.3%; Pred. No. 0; Mismatches 1269; Indels 51; Gaps 5;			
Matches 2177; Conservative 0;			
QY	162	caggttgaaccctgtattatgttcggataagtggttcgtgcctgatttgygtacataca	221
Db	760	caggtatgcctattatcaagtgaattggtgctgggtgctggttcgtctgtacatca	819
QY	222	gatttccctgtgattatataactgcagcacagacagcaagagagattcgaaaacagtttt	281
Db	820	ggtttcatttggtgcttcgagcgcgcggaagacaataacaaattgaaaacagtttt	879
QY	282	tcatcagtttggcacagacatcgctggtttatagctgtgacatcggtgaaactaa	341
Db	880	tcatgctataaagcagacagagatagctggtttgatgtgcagcagatgtggyggaactaa	939
QY	342	cactgcagacagacatgtgac---aaatcagtgatgttatgtgagataaagattgcct	398
Db	940	caaccgacctacagagatgctcttaagaattcaacagttatgtgtgaaaaatttgaat	999
QY	399	gtgtgttcaaaaacatgctacttttcgatttgccctgagctgagttgttggtaagggctg	458
Db	1000	gtcttcttcagtataagcaaatcttttcactggtttatagtaggatttaacacgtgctg	1059
QY	459	gaactacacctagtgacttatccaagctctcccttataatagcttcaagcgacagatg	518
Db	1060	gaagctaacaccttgatatttggcaccacagctcgttcttcttgacatgctgaagctgcgtctg	1119

QY	519	ttctaggatggtcattcattctatgacagatgaaggaaattaaagtccattatccaagctggygc	578
Db	1120	gccaagaatctactatctcattctactgataaagaaccttagcgtatgcaaaagcttggagc	1179
QY	579	tgtggcagaagaagctctgttcaatcccgagcaagctacagcttagccttagggccgaaggaa	638
Db	1180	agtaactgaaggagctcttgcgcagcaattgaactgtgatgtcatttgggggacaagaagaa	1239
QY	639	agaattctcaaaagttatcacagaatctccaagaatgcaaaagatlttggcataaaagagac	698
Db	1240	agaacttggaaaggtacaacaaaatttagaagaagtaaaagattgggataaagaagc	1299
QY	699	tatagtctcaaaagtgctccttggtgtgtgtgtatcttctaigaatggaaacctatygact	758
Db	1300	tattcaagccaatttctatagtggtgtgtcttccgctcgatctatgactcttaatgtctc	1359
QY	759	tgccttttgatgtagaaccttcctgtatcttaatgtagaacctgtgatatcacctcggac	818
Db	1360	ggcctctgtgatacggacaaccttgcctctccaaggga-----atatctattggaca	1413
QY	819	tgtctcgtgttcttctcttagtgaatccataagcagttatgtgattggagccaactcc	878
Db	1414	agtaactcactgtatcttcttctgtartaattgggctttagtggttggacagcatctcc	1473
QY	879	tcaatttgaacctctgcataagaccggaagagctgccttcaatatlttccaagttatga	938
Db	1474	aagcatctgaagcaatttgcataatgcaagagagagcattatgaactctccaagataatga	1533
QY	939	taagaaccacaagatagataactttccacaagctgatatataaacctgcgaatccataagag	998
Db	1534	tactaagccaagatctgacagcgcattctcgaaaggttggcacaacaaccgatatattaaagg	1593
QY	999	aaactgtggaatttaaaaatgttcttcttaattatccatacgaaccatctacaagattct	1058
Db	1594	aaatttggaaatccaagaatgttcaactcagttacccaatccgaagaagaagttaaagattct	1653
QY	1059	gaaaagctctgaatcctcagaattaaagtcctggagagacagtcgccttggtcgtctcaatgg	1118
Db	1654	gaagggccttbaaacctgaaggtgcagagatgtggcagagcgttgccctggttggaaaacaattg	1713
QY	1119	cagtgaggaaagatgacgttagtccagctctgcgcagaggttatgatcccggaatgaagctt	1178
Db	1714	ctgtgggaaagaagcaacaacatccagcgtgacgtgacagagctctctatgaccccaagaggggat	1773
QY	1179	tatcactgtgtgagatgagatatacactcagagcttcaaaatgtgcggatatacgagaccat	1238
Db	1774	ggtcagatgtgtgagagagagatatattagaccataaattgtaagtgttctaagggaaatcat	1833
QY	1239	tggagtggttagtcaagaagcctgtttgttcggygacacacatcagttacaatatcaagta	1298
Db	1834	tgggtgtgtgtagcagaaacctgtatgttctgcacacacatcagcttgaaaacattcgtcta	1893
QY	1299	tggacagagatgagtgatctgataaagagatggagaagagcgagcaagggaagaagaatgtgcta	1358
Db	1894	tggcgtgaaaatgltccacatgtgatgagatgagaaagctgttccaaggaagccaattgccta	1953
QY	1359	tggattatcatatgaggttccctaaataaattatatacttggtaaggggaaaagaagctca	1418
Db	1954	tgaacttatcatgaactgcctcataaatttgacaacctgtgttgtagagagaaggggccca	2013
QY	1419	aattgagtggaaggcagaacaagagagatccgaactgtctgtgctttagttcgaaaacccaa	1478
Db	2014	gttgagtggtgggagagaagcagagatgcacatgcacgtgcctgtgttcgcaaaccccaa	2073
QY	1479	gattctgattttagatgtagcgtacgctctgccttgattcagaagaagcgaagtacgctgttca	1538
Db	2074	gattcctctctgtgatatgaggtccacgtttagccttggacaacagaagcgaagaagtggttca	2133
QY	1539	agctgcacatggagaagcggagcaaaagttcggaactacaaatcgttggtagacaacgaattcc	1598
Db	2134	ggtgtgctctgtgataaggccaaaagaagtcggacaaccaactgtgtgatagctcatcgtttgtc	2199

QY	1599	tactcttcgaagctgcagatttgatttgatcttgacccctaaagaatggaatgctgtgcggagaagg	1658
Db	2194	tacagcttgtaatgtctgacgcacacgcgtggtcttcgatagtatgagctatctgttcggaaaag	2253
QY	1659	agcacatctgaactaaatgagcaaaagaggctatcatctacactgtgtgatcagagga	1718
Db	2254	aaatcaatgataactcctacagaaaagaaagagcatttacttcaactctgttcacaatgcagac	2313
QY	1719	tattaaaaaagctgatatgacacagatgagatcatalgtatctc-----1762	
Db	2314	agcaggaatgaagtgtgatttagaaaatgacgctgatcatccaaaagtgaattgtgc	2373
QY	1763	-ctgaagaagaacacacatcactctcctctgacctctgtgaagagcalca-----1810	
Db	2374	cttggaaatgttcttcaaatgataltcaagaatccagcttcaataagaaaagatcactgtag	2433
QY	1811	-----agtcagacttcattgtaacaagctgcgaggaatcccaatccaaatctaaagat	1859
Db	2434	gaggtccgtgtgatacacaagcccaaggaagaagcttgatctccaagaagagctcgtgatga	2493
QY	1860	aagctcttcgtaagctctctctatlttaaaaatttaagtlttaacaagaagcctgtaatggcctt	1919
Db	2494	aagataactccagttccctcttggagattatgagtaagctaaatttaactgatgacctta	2553
QY	1920	tgatgtctcgtggacattggcctctgtctctaaatgaaacgtttcatccagatatttccat	1979
Db	2554	ttttgtgtgtgtgtatlttggtccattataatgtgaaggctgcgaaccaagcagattgtcaat	2613
QY	1980	catctttgaaaaatatcaacatgttt---ggaatatgatataaacacatcaataagca	2036
Db	2614	aactatttcaaaagatataggggtctttacaagaattgatcctcgtgaacaaaaagaca	2673
QY	2037	tgatcgagaaatattatccatgatataatccgcatltttggtgtatattgtctgttcagtta	2096
Db	2674	gaatagtaactctgttttcaactatgtttcttagcctcctggaaattattctttattacatt	2733
QY	2097	ttcatatgcaggaattattttacgcgacagcagggaaatttaacgatgagataagaaca	2156
Db	2734	ttctcttcaggggtttacaatttcggaagcctgtgaagatccctcaccaagcggctcgcgata	2793
QY	2157	ctttgacctcaagcagatlttatcatcagatattgcctgtgtttgtatgtaaaaggaanaag	2216
Db	2794	catggtcttcgcgatccaatgctcagaagaagatgttggtgtgtgttgatgcacctaaaaaac	2853
QY	2217	caacagagagctgtgacaacaataatgacccaatagataatgacaaatccaagagacaaag	2276
Db	2854	caactgagacattgacatcacccagcgcctgcgcacatgatgtcgtcaagttaaagggcctatag	2913
QY	2277	ttccagagattgagcgtcttaacacaaaaatgcacatacattacatgggacttcaatcatcttc	2336
Db	2914	ttccagagcttgatlaattaccacagatatlaagcaaatctgtggacaggaataatlatatc	2973
QY	2337	ctttatatatgtagtgagatgatacatctcctgtgtctcgagtatgtgtccagttacgttcgcgt	2396
Db	2974	cttcatcctaaigtgtgcaactaacctacgttactctcttagacaaattgttaccatcatgtgcaat	3033
QY	2397	gacacgaatgatatgaacccgcagccaatgtgattgttcacaaacagaataagcaagaact	2456
Db	3034	agcagagagatgttgtaaatgaanaatgttgcctgacaaagcagtaagaataagaagaact	3093
QY	2457	taagcatgctggaagaatgacacatgaaagccttgtagaataatagctactatagtgtcatt	2516
Db	3094	agaagatgctgggaagaatgcctactctgaagaacataagaaaactctcgaaacccgtgttctctt	3153
QY	2517	aacaagaagaagaagccttcgagcaaatgtatgaaggatcgttcagactccaacagaa	2576
Db	3154	gactcaggaacagaagaatttgaacatatgtatgctcagagtttgcaggtacacataacaga	3213
QY	2577	tactctgaagaagaagacagattatggaagctgttatgatatcaagcagcttlatata	2636
Db	3214	ctctcttgagaagaagcacacactcttggatatacttctccttcaaccagcgaatgatgta	3273
QY	2637	ttttgccaatgcagcaggggtttcagatttggagcctatlaattcaagctgcagcaatgtac	2696

Db 3274 ttttccatgtctgtagtcttccggttggagccctacttgggggacataaactatgag 3333  
Oy 2657 cccagaggagcatgtcatagtttactgtcaattgcatatgagcatgagccatgga 2756  
Db 3334 ctttgaagatgtcttctgtagtatttccagctgttcttggtagcagtcggtgg 3393  
Oy 2757 aacgtctgttggcttgaatattccaaagccaaatcggtgggtgtgcatctgttc 2816  
Db 3334 agtcagttcatcttccctgtagtaccagaaagccaaatattccagccatcat 3453  
Oy 2817 ctgttggaaaagaaaccataatagacagccgacgaaggaagaaagaaagacac 2876  
Db 3454 gatcatlgaanaaaaccccttggattgacagctacagcaaggaagccatgacga 3513  
Oy 2877 atgtgaaggaattagagcttgcgaagctcttcttctatcatatgcgccagatg 2936  
Db 3514 atgtgaaggaattagagcttgcgaagcttgcgaagcttgcgaagcttgcgaag 3573  
Oy 2937 ttcatctcgttgcctatccctcagatgagcgaggaaggaagacagatcttgg 2996  
Db 3574 cccagtgctcagagacccgagcctggaggtgaagagggccagagctgtgtgtg 3633  
Oy 2997 gagcagcgctgtggaaaagacacttcttcaacttctgcagagactttagacccgt 3056  
Db 3634 cagcagtgctgtgtggaggaagacagctgttccagcttctgcagcggttctacga 3693  
Oy 3057 gcaaggaacagtgctgttgaatgtgtgtgaatgaagaaatgaatgaatgaatg 3116  
Db 3694 ggcaggaagagtgctgttgaatgtgtgtgaatgaagaaatgaatgaatgaatg 3753  
Oy 3117 ttcccaatagcaatctgttctcagaagcctgtgtcttcaacttgcagcatgtgtga 3176  
Db 3754 agcacactgt 3813  
Oy 3177 catgcctatgtgtagcaacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3236  
Db 3814 catgtcctatgtgtagcaacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3873  
Oy 3237 tgcagcaaatatcatcttatttgaaggttctcctgagaataatacaacagatgg 3296  
Db 3874 ggaagccacacatatactgtccttcatcgaatcgaatcgaatcgaatcgaatcga 3933  
Oy 3297 actgaaggaagacagcttcttgcggcgcaagaacaaagacatgatttgcgaagct 3356  
Db 3934 agacaaaggaactcagctctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3993  
Oy 3357 tctccaaaacccaaatttatttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3416  
Db 3994 tgttgaagagcctcatatttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4053  
Oy 3417 tgaagaagtgtgtcagcatcttcttgaataagccaggaaggaaggaatgtgtgt 3476  
Db 4054 tgaagaagtgtgtcagcatcttcttgaataagccaggaaggaaggaatgtgtgt 4113  
Oy 3477 caactacagcctcttgcatttcaataagcaagcaagcaatgtgtgtgtgtgtgt 3536  
Db 4114 tgcataccgctgtgtcaccatcccaaatgtgtgtgtgtgtgtgtgtgtgtgtgt 4173  
Oy 3537 gataaaggaaggaatcattcagaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3596  
Db 4174 agtcaaggaagcattgcaagcattcagaagcgtgtgtgtgtgtgtgtgtgtgtgt 4233  
Oy 3597 agtgaatgcacagtcag 3613  
Db 4234 ggtcagtgctcagcgtg 4250

RESULT 5  
US-08-583-276-18  
; Sequence 18, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:

APPLICANT: McDonald, Kevin T.  
APPLICANT: Niemhuis, Arthur  
APPLICANT: Tolstoshev, Paul  
TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carella, Byrne, Bain, Gillfillan,  
ADDRESSEE: Cecchi & Stewart  
STREET: 6 Becker Farm Road  
CITY: Roseland  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch diskette  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: DNA V2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/583,276  
FILING DATE: 05-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/332,444  
FILING DATE: 31-OCT-1994  
APPLICATION NUMBER: 07/887,712  
FILING DATE: 22-MAY-1992  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4669 bases  
TYPE: nucleic acid  
STRANDEDNESS: singular  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: Genomic DNA  
US-08-583-276-18

Query Match 36.1%; Score 1308.2; DB 2; Length 4669;  
Best Local Similarity 62.1%; Pred. No. 0;  
Matches 2173; Conservative 0; Mismatches 1273; Indels 51; Gaps 5;  
Oy 162 caggttaccctgtatattgttgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 221  
Db 760 caggttaccctgtatattgttgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 819  
Oy 222 gatttccctgtgtatatttgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 281  
Db 820 gatttccctgtgtatatttgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 879  
Oy 282 tcaatcagtttggcagacacatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 341  
Db 880 tcaatcagtttggcagacacatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 939  
Oy 342 caatcagtttggcagacacatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 398  
Db 940 caatcagtttggcagacacatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 999  
Oy 399 gttgttcaaaatgtcttcttgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 458  
Db 1000 gttgttcaaaatgtcttcttgaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1059  
Oy 459 gaaatcaccctagttacatcagcttcccttataatgtgtgtgtgtgtgtgtgtgtgt 518  
Db 1060 gaaatcaccctagttacatcagcttcccttataatgtgtgtgtgtgtgtgtgtgtgt 1119  
Oy 519 ttctagatgtgtatcatttgcagcagtaaggaatgaatgtgtgtgtgtgtgtgtgtgt 578  
Db 1120 ttctagatgtgtatcatttgcagcagtaaggaatgaatgtgtgtgtgtgtgtgtgtgt 1179  
Oy 579 tgtgtgagaagaaggttcttgcattcattcagcagcagcagcagcagcagcagcagc 638

Db 1180 AGAGCTGGAAGAGTCTTGGCAGCAATTAGAACGTGATTGCAATTTGGAGACAAAAAGAA 1239  
QY agagactccaaggatatacaagaatctcaaaagtcaaaagatttggcataaaagac 698  
Db 1240 AGAAGCTTGAAGGTACACAAAAATTAGAAAGCTTAAAGAAATTTGGATTAAGAAAGC 1239  
QY tatagctcaaaaagtgctcttggtgctgctgcttcttataatgaatgaatgaactatgact 758  
Db 1300 TATTACAGCAATATTCTATAGTGTGCTTCTCTGCTGATCATCATCTTATGCTCT 1359  
QY 759 tgcgttttgatagaaacctcttgatcttcaatggaacctggaataccatcgagac 818  
Db 1360 GGCCCTTCTGATGGGACCACTTGGTCTCTCAGGGGA-----ATATTCTATTGGACA 1413  
QY 819 tgccttgcgtttcttctttagtgaatccatagcagttatgcatgagcagagcc 878  
Db 1414 AGTACTACGTATCTTCTCTGATTATTTGGGCTTTTATGTTGACAGGCATCTCC 1473  
QY 879 tcaacttgaaccttcgaataagccgaagagctgccttcataatctccaagtattga 938  
Db 1474 AAGCATTTGAAGCTTTTCCAAATGCAAGAGAGCAGCTTATGAATCTTCAAGATTAATGA 1533  
QY 939 taagaaacccagatagataaactttccacagctggaataaactgaatccatagaag 998  
Db 1534 TAATTAAGCCAAAGTATGACAGCTATTCGAAGAGTGGGCACAAACCATATATTAAGGG 1593  
QY 999 aacttgggaatttaaaatttctcttcaatatacatcaagaccatcaatcaagattct 1058  
Db 1594 AATTTTGAATTCAGAAATGTTCACTTCACTTACCTTACCCATTCGAAAGAAATTAAGATCTT 1653  
QY 1059 gaaagctcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaat 1118  
Db 1654 GAAGGCGCTGAACCTGAAAGTGACAGAGTGAGGACAGCTCCCTCCCTGGTGGAAACAGTGG 1713  
QY 1119 cagtggaagaatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1178  
Db 1714 CTGGGGAAGAGCAACAGCTGATGATGACAGAGCTCTATGATACCCACAGAGGGGAT 1773  
QY 1179 tatcatgttgatgagatgagatgagatgagatgagatgagatgagatgagatgagatgag 1238  
Db 1774 GGTGAGTGTTCATGAGACAGATATTAAGACCATTAATGATGATTTACAGGGAATCAT 1833  
QY 1239 tgaagtggttagtcaagaagccgttttgctgaggaacccaatcagtaataatcaagta 1298  
Db 1834 TGGTGTGTGATGAGAGACCTGATTTGTTGCCACACGATGACTGAATTCGCTA 1893  
QY 1299 tgaagcagaatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1358  
Db 1894 TGGCCGTGAAGATGTCACATGATGATGATGATGATGATGATGATGATGATGATGATG 1953  
QY 1359 tgatattatcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1418  
Db 1954 TGACTTTATCTGAAACTGCTCTCAATAATTTTGACACCTGCTTGGAGAGAGAGGGCCCA 2013  
QY 1419 aatgaagtgaaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1478  
Db 2014 GTTGAGTGTGGGAGAGAGAGATGCGCATTTGACAGTCCCTGTTGCCAACCCCAA 2073  
QY 1479 gatctgaattttagatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1538  
Db 2074 GATCTCTCTGCTGATGAGGACGCTGACGCTTGGACACAAAGCGAAGAGTGGTTCA 2133  
QY 1539 agctgaacttgaag 1598  
Db 2134 GGTGGCTCTGATTAAGCCAGAAAAAGTCCGACACCATTTGTGATAGCTCATCGTTTGT 2193  
QY 1599 tactattcgaagtgagagatttgatgttgagtgagtgagtgagtgagtgagtgagtgag 1658  
Db 2194 TACAGTTGCTGAATGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2253  
QY 1659 agcaaatgctgaactaatgagcaaaagagtgctatatataatcaacttgatgagtgagga 1718

Db 2254 AATCATGATGAACTCATGAAAGGAAGGCAATTTACTTCAAACTTGTCAAAATGCGAGC 2313  
QY 1719 tatataaaagctgagtgaa-----cagatggagtgcaatgacatattctactg 1765  
Db 2214 AGCAGAAATGAAGTGAATTAAGAAATGCAAGCTGATGAATTCAAAAGTGAATTTGATGC 2373  
QY 1766 aaagaagaaccaactcaactcctctctgacactctgagagagatcaagtcaactctg 1825  
Db 2374 CTTGAAGATGCTTCAATGATTTCAAGATCAGCTAATAAGAAAAGATCACTGCTAG 2433  
QY 1826 acaagtgtaggaatccaccc-----aatcaagaagat 1859  
Db 2434 GAGTGTCCGTGGATCACACCCAGACAGAACTTAGTACCAAAAGGCTCTGGATGA 2493  
QY 1860 aagcttcctgaagtgctctctatataaaatttlaaagtltaaacagcctgaatgacctt 1919  
Db 2494 AAGTATACCTCCACTTTCCTTTGGAGATTAAGACCTAAATTAATGAAATGGCCTTA 2553  
QY 1920 tgggtctgaggagacattgctctgcttctaataatgaaatgcttcaatcagatattccat 1979  
Db 2554 TTTTGTGTGGTGTATTTTGTGCAATTAATGAGAGCCCTCGAACCCAGCATTTGCAAT 2613  
QY 1980 catcttgcaaaaattataacacagttt---ggaataatgataaaaccaattaaagca 2036  
Db 2614 AATATTTCAAGATTTATAGGGCTTTTACAGAAATGATGATCTGAAACAAACGACA 2673  
QY 2037 tgaatgcagaatattatccatgatalatcgtatcttgggtgagtgatgcttgctgacgtta 2096  
Db 2674 GAATATCAACTGTTGTTACATATTGTTCTAGCCCTTGGATTAATTTCTTTATTAATTT 2733  
QY 2097 ttcaatgagagatlatlttccgagagagagagagagagagagagagagagagagagagag 2156  
Db 2734 TTTTCTTCAAGGTTTACATTTTGGCAAAAGCTGAGAGATCTCCACCAAGCGCTCCGTA 2793  
QY 2157 ctggccttcaagcagatgatalatcagagatgagcctggttggatgagagagagagagag 2216  
Db 2794 CATGTTTTCGATTCATGCTGACAGAGATGATGATGATGATGATGATGATGATGATGATG 2853  
QY 2217 cacaagagcttgaacaacatattagccaatagatatagcacaatltcaaggaagcaagag 2276  
Db 2854 CACTGAGACATTGACTACAGAGCTGCGCAATGATGCTCTCAAGTTAAAGGCGCTATAGG 2913  
QY 2277 ttcaagagtggtgttctaacaacaaatgcaactcaactcagtgagcttcaagtgatcattc 2336  
Db 2914 TTTCCAGGCTTCTGTAATTTACCCAGAAATATGCAATTTTGGACAGGAATTAATTAATC 2973  
QY 2337 ctltatataatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2396  
Db 2974 CTTCATCTATGTTGGCAACTTAACCTCTTACTCTTACCAATTTGACCCATCATTTGCAT 3033  
QY 2397 gacaggaatgataatgaacccagcaatgagtgagtgagtgagtgagtgagtgagtgagtg 2456  
Db 3034 AGCAGGAGCTGTGTAATGAAGAAATGTTGTGACAAACACATGAAGATTAAGAAAGACT 3093  
QY 2457 taagcatgctgagaaatagcaactggaagcttggagagataatgattactatagtgact 2516  
Db 3094 AGAAGTCTCTGGGAAGATCGTACTAGAGCAATTAATAAACTTCGAAACGTTGTTCTTT 3153  
QY 2517 aacaagagaaagccctcgagcaaatgataatgaagaga tgcctcagactcaacagaaa 2576  
Db 3154 GACTCAGAGAGAGAAAGTTTGAACATATGATGCTAGAGGTTGGAGGATACATACAGAAA 3213  
QY 2577 tactcgaagaagacacagatatttgaagatgataatgacatcagcaatgccttatata 2636  
Db 3214 CTCTTTGAGGAAGACACATCTTTGGAATTAATTTCTTTCAACCCAGGCAATGATGTA 3273  
QY 2637 ttctgccaatgagcagaggttctgagtgagtgagtgagtgagtgagtgagtgagtgagtg 2696  
Db 3274 TTTTTCCTATCTGATGTTTCCGTTGGAGCCCTACTTGGGCACTAAATCTATAGAG 3333  
QY 2697 cccaagagagcagtgatcagtgatatttcaatgcaatgacatgagtgagtgagtgagtgag 2756  
Db 3334 CTTTGAGATGTTCTGTTAGTATTTTCAAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3393







QY 579 tctgcaagaagctctgtcatcaatccgaacagatagcctttagggcccaagaa 638  
1180 AGTAGCTGAAGAGCTTGGACCAATTAGATGATTTGAGAGCAAAAGAA 1239  
QY 639 agaatctcaagatatacaagaatctcaagaatgcaagatttggcctaaagagac 698  
1240 AGAAGCTGAAGAGTACAAACAAAATTTAGAAAGCTAAAGAAATGGGATTAAGCAAGC 1299  
QY 699 tataagctcaaaaagtgctctgtgctgtgactcttcaatgaaactatgact 758  
1300 TATTACAGCCAAATTTCTATAGTGCTGCTTCTCTGATCATGATTTATGCGT 1359  
QY 759 tgccttttggatagaaacctctgtatcttaatagtagaacctgataaccacggac 818  
1360 GGCCCTCTGATGGAGCCACCTGTGCTCTCTCAGGGGA-----AATTTCTATTGGACA 1413  
QY 819 tgtctctgctgttctctttagtgaatcaatagcaagtattgcatgtgagacagctccc 878  
1414 AGTACTCACTGATTTCTTTCTGATTAATTTGGGCTTTTATGATGGACAGCATCTCC 1473  
QY 879 tcaacttgaaccttcgcaatagcccgagagctgaccttcatatttccaggttatgca 938  
1474 AAGCATTTGAAGCATTTTGCMAATGCAAGAGAGCAGCTTATGAATCTTCAAGATTAATGA 1533  
QY 939 taagaacacccagatagataaacttccacagctgagatataaactgaaatccatagaag 998  
1534 TAAATTAACCAAGTATTCAGACCTATTGCAAGAGTGGGACAAACAGATTAATTAAGGG 1593  
QY 999 aactgtagaatttaaaatgcttcttcaatctatccatcaagacacatctataagaattct 1058  
1594 AAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAANAAGATTAAAGATCTT 1653  
QY 1059 gaaagctcgaatctcgaatlaagctcgaagagacagctcctgtgctgctcctcaatgg 1118  
1654 GAAAGGCTGGAACCTGGAAGTGCAGAGTGGGACAGGCTGCTGCTGTTGAAGACAGTGG 1713  
QY 1119 caatggaagaagtaagctgtagctcagctctcgcagaggtatagatccgagatgagctt 1178  
1714 CTGCGGGAAGAGCAACAGTCCAGCTGATGACAGAGCTCTATGACCCACAGAGGGGAT 1773  
QY 1179 tataatggttagtaagaatgacatcaagaacttaaatgtgagctatctgaagacatat 1238  
1774 GGTCAGTGTATGACAGATTAATGACCATTAATGATTAAGTTTCTACGGGAATGAT 1833  
QY 1239 tggagtggttagtaagaagcctgttctgtcgtggaacccacatcaglaaacaatcaatga 1298  
1834 TGGTGTGTGATGACAGCACTGATTTGTTGCCACAGATAGCTGAAACATTCGCTA 1893  
QY 1299 tggagcgaagatgactgtagctatgaagagatgagagagcagaaggaacaatgtgta 1358  
1894 TGGCCGTGAAAATGTCAACATGATGAGATGAGAAAGCTGTCAAGGAAGCCAAATGCTA 1953  
QY 1359 tgatttatcaatgtaggtctcctaataatlaatacatgtgtagaggaaggaagagctca 1418  
1954 TGACTTTATCTGAAACTGCTCATTAATTTGACACCTGTGTGAGAGAGAGGGGCCCA 2013  
QY 1419 aatgctggaaggaagaagaacagatcgaactgctgctccttaagctgaagacccca 1478  
2014 GTTAGAGTGTGGGAGAGAGAGAGATGCGCATTCAGCTGCCCTGTTCCGAACCCAA 2073  
QY 1479 gattctgatttagatgagctacgctcgtccgtgattcagaagaagaagcagctgtca 1538  
2074 GATCTCTCTGCTGATGAGGCGACGTCAGCTTGAGACACAAAGCAAGAGTGTGTTCA 2133  
QY 1539 agctgcaactggaagaaggaacaaagtctggaactacaactgtgtgtagcaccacgacttc 1598  
2134 GGTGGCTCTGATTAAGGCAAAAGGTGGAACCACTTGTGATGATGCTCATCTGTTTGC 2193  
QY 1599 taactatggaagtgagattgtagtctgacccctaaagaatgaaagctgagcggaagaa 1658  
2194 TACAGTTGTAATGCTGACGTCATGCTGTTGATGATGAGATGATTTGAGAAAGG 2253  
QY 1659 agcaaatgctgactaatgagcaaaacgagcttatattatcaacttgtgtagtcacagga 1718

DB 2254 AAATCATGATTAACCTCAAGAAAGAAAGGCAATTTACTTCAACTGTGCATATGCAAGAC 2313  
1719 tattaanaagctgtagacagatgagatgacatctcta----- 1762  
DB 2314 AGCAGGAATATGAAGTTGAAATAGAAAATGCAAGCTGATCAATCCAAAAGTGAATGTATGC 2373  
1763 -ctgaagaagaagaccatcactctcctcgtcagctcgtgaaagacatca----- 1810  
DB 2374 CTGGAATGTCTTCAAAATGATTAAGATCCAGTCTATTAAGAAAAGATCACTCGTAG 2433  
1811 -----agtcagaactcattgacaagctcgaggaatccaccccatctaaagat 1859  
DB 2434 GAGTGTCCGTGATCACAGGCCAACAGAAAGCTTAATGTAACCAAGAGCTCTGGATGA 2493  
1860 aagctctcgaagctcctcattaaatlaaagtltaaaagaacgctgtagtgcctt 1919  
2494 AAGTATACCTCAAGTTTCCTTTGGAGATTAATGAAAGTAAATTAAGTGAATGGCTTA 2553  
QY 1920 tgtgtctgggacatgagctcgtctcctaagtgaactgtlcaacagatctccat 1979  
2554 TTTTGTGTGTGATTTTGTGCATTAATTAAGAGAGCCCTGCMAACGACATTTGCAAT 2613  
QY 1980 catctttgcaaaaatataacatgltt---gaaatataataaaccacattaaaga 2036  
2614 AATATTTTCAAGATTAAGGGGTTTATCAAGAAATGATGATCTGTAACCAAAACGACA 2673  
QY 2037 tgaagcagaatattatccatgatatctgcatatttgggtgtatttctgtcgtca 2096  
2674 GAATGTAACTGTTTTCATATCTTTCTAGCCCTTGGAATTAATTTCTTTATTAATTT 2733  
DB 2734 TTTCTCTAGGTTTACATTTTGGCAAGCTGGAAGATCTCTACCAAGCGGCTCCGATA 2793  
QY 2157 ctggcctcacaagcagatgatatcaagatatgcccgtgttgtagaagaagaaagacag 2216  
2794 CATGTTTTCGATCCATGCTCAGACAGAGATGAGTTGTTTCATGACCTTAATAACAC 2853  
QY 2217 caaagagcctgacaacaatatlaagcatalagataagacaaatccaaaggaagacag 2276  
2854 CACTGACATTTGATACAGAGCTGCGCAATGATCTCTCAAGTTAAAGGGCTATAGG 2913  
QY 2277 ttccagagctggcgtcttaacaaatgcaactaacaatggaacttcaagtatattc 2336  
2914 TTCCAGGCTTGGCTGTAATTAACCCAGAAATATAGCAATTTGGACAGAAATTAATATC 2973  
QY 2337 cttaataatgagtgagatgacatctcgaatcttgcagatgttgcagatgagctg 2396  
2974 CTTCATGTATGTTGGCAATACACTGTTACTTTACCAATTTGATCCATCATTTGCAT 3033  
QY 2397 gacaggaatgattgaacccgacagcaatgactgatttccacaacaagaatgaagaaact 2456  
3034 AGCAGAGCTGTGTAATTAAGAAATGTTTGTGACACACAGCAAGAAATGAAGAAAGACT 3093  
QY 2457 taagcagctggaagaatagcaactgaaagcttggagatalaogtactatgvtcaat 2516  
3094 AGAAGTGTCTGGAGATGCTGTAAGCAATTAAGAAATTTCCCAACCGTGTCTTCTT 3153  
QY 2517 aacaagggaaaaagccttcgagcaatgtatgaagaagtgctcgaactcaacaagaaa 2576  
3154 GACTCAGAGCAAGAAATTTGAACATATGATGCTCAGAGTTTGACAGTACATTAAGAAA 3213  
QY 2577 taacctgagaagaagacagatattggaagctgctatgcatcaagcagctcattata 2636  
3214 CTCTTTGAGGAAGACACATCTTTGGAATTAATTTCTCTCACCCAGGCAATGATGA 3273  
QY 2637 ttctgcatagcagaaggttgcattgtagagcatttaattcaagctgtagcagatgac 2696  
3274 TTTTTCATATGCTGATGATGTTCCGGTTTGGAGCCTTACTGTGGGACATTAACATCATG 3333  
QY 2697 cccagaggcagatgctcaatgatttcaactgcaatgacatagagctatgagctatgagaa 2756

```

DB 3334 CTTTGAGGATGTTCTGTAGTATTTTTCAGCTGTGTTGTTGGTCCATGGCCGTGGGGCA 3393
QY 2757 aacgctcgttttggctcgtgaatattccaaagccaaatcgggggcgtgcacatcgtttc 2816
DB 3394 AGTCAGATTCTTGTCTCCGACTATGCCAAAGCCAAATATCAGCCACCATCATCAT 3453
QY 2817 ctgttggaaagaaacaaatattagacagccgagtcgaagaaagaaagaaagcagacac 2876
DB 3454 GATCATATTGAAAAAACCCTTTGATTGACAGCTACACAGCAAGGAGCTTAATGCGGAAC 3513
QY 2877 atgtgaaggaatattagatcttcgaagaagtcctctctatccatctgcgcgaatgt 2936
DB 3514 ATTGGAAGGAANTGTCACATTTGCTGAAGTGTGATTTCACTATCCACCGGACGACAT 3573
QY 2937 ttcatcctcgttggtctatccctcagatltgaagcgaagaaagacagtcagatlttgg 2996
DB 3574 CCCAGTGTTCAGGGACAGTACGAGTGAAGGAGGCGACAGCCCTGCTGTGTGG 3633
QY 2997 gaggcagcgttgggaaagacactctgttcaactcttcagagagacttataccggt 3056
DB 3634 CAGCAGTGTGCTGGGAGAGACACAGTGTCCAGCTCTGGAGCGGTTCTACGACCCCT 3693
QY 3057 gcaaggaacagtgctgttgaatgtgtgaatgaagaatgaatgaatgaatgaatgaatga 3116
DB 3694 GCGAGGGAAGAGTCTCTGATGTCGAAGAAGAAATGAAGCCACTGAATGTTCAAGTGTCC 3753
QY 3117 ttcccaatagaaatcgtcttccctcaagagcctgtgtcttcaactgaagatgtcgtga 3176
DB 3754 ACACACACCTGGGACATGCTGCCAGAGCCCATCTGTTGATGACGAGCATGCTGTGAGAA 3813
QY 3177 catcgcctatgttgaacaaagcgtgtgtgtccatagatgaatcaagaacgcga 3236
DB 3814 CATTGCTTGTGAGAACAGCGGGTGTGTCAGGAAGAGATGCTGAGGGCAGCAAA 3873
QY 3237 tgcgaagaatccatctcttatttgaaggtctccctgaagaatatacaacacaagttg 3296
DB 3874 GAGGCGCAACATCATGCTCTTCATGACAGTCAGTCTTAATAATATAGCTAAAGTAGG 3333
QY 3297 actgaaggaagacagacttcttgcgcccagaaacaaagactatgttcaagggctt 3356
DB 3934 AGACAAGGAACACTCAGCTCTGTGTGGCCAGAAACAGCATGTCATACCTGTGCGCT 3393
QY 3357 tctccaaaacccaaatttatttatttgaatgagtgagcactcagccctgataatga 3416
DB 3994 TGTTAACACGCTCATATTGCTTTGATGTAAGCCACGTCAGTCTGTGATACAGAA 4053
QY 3417 tgaagaagtgttgcagatgccttgataaagccaggaacggaagacatgtcgttga 3476
DB 4054 TGAAGAAGTGTGTCAGAGAGCCCTGGAGCAAGGCAAGGAGGCGCCGACCTGCAATGTGAT 4113
QY 3477 cactcaagcgtctctgcaatctcagaacgcaagattgatagtgttctgcacaatgga 3536
DB 4114 TGCTCACCGCTCTCCACCATCCAGATGAGACTTAATAGTGTGTTGATGATGAGAC 4173
QY 3537 gataaaggaacaaagacatcataagagctcccgagaatccgagacataatattga 3596
DB 4174 AGTCAAGGAGACATGGACAGGATCAGAGCTGTGGCACAGAAAGGCAATCAATTTTCAAT 4233
QY 3597 agtgaatgcagatcag 3613
DB 4234 GGTGAGTGTCCAGGCTG 4250

```

```

RESULT 7
US-08-784-649A-1
: Sequence 1, Application US/08784649A
: Patent No. 5830697
: GENERAL INFORMATION:
: APPLICANT: SIKIC, Branimir I
: APPLICANT: Chen, Gang
: TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
: TITLE OF INVENTION: CYCLOSPORIN MODULATION
: NUMBER OF SEQUENCES: 5

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg. No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-1

```

```

Query Match 36.0%; Score 1305.2; DB 2; Length 4264;
Best Local Similarity 61.9%; Pred. No. 0;
Matches 2163; Conservative 0; Mismatches 1283; Indels 48; Gaps 4;

```

```

QY 162 caggttgaccctgtatatttggaaatagtggtgcgcttgatttggtaacataca 221
DB 477 CAGGTATGCTATTAATTAACGTGGAATGTTGCTGGGCTGTGCTGCTCATCAATCA 536
QY 222 gattcctgttgatatactgcagcagcagcagcagcagcagcagcagcagcagcagc 281
DB 537 GATTTCATTTTGGTGGCTGGCAGCTGGAAGACAAATACAAATATGAAGAAGCTTTT 596
QY 282 tcatcagtttggcagacagacatcggcgtgttgatagcttgcgttgaaactaa 341
DB 597 TCATGCTATATGCGACAGAGATAGGCTGTTGATGTGCAGAGTGTGGGAGCTTAA 656
QY 342 cactcgatcagcagacatgcacaatcagtgatgtatgtggagataagatgtctcgt 401
DB 657 CACCCGACTTACAGATATGTCCTCAAGATTAAAGAAATGTCACAAATTTGCAAT 716
QY 402 gttcaaacatgctctacttcttgcattggtcgtgcgttggttggatgaagcgtgaa 461
DB 717 GTTTCATCATATGCAACATTTTTCACCTGGTTTATAGTAGATTTACAGTGGTTGAA 776
QY 462 actcaccctagactatccacgctcctctataatggtctcagcggcagcagcttgc 521
DB 777 GCTAACCTTGTGATTTTGGCCATCACTGCTGTTGACTGTGACACTGCTGCTGGCG 836
QY 522 taggatgtcatctcatatgcagtaagaaatgaagcctatccaaagctgggctgtc 581
DB 837 AAAGATCTACTCTTCATTTACTGATTAAGAAACCTTATGCTATGCAAAAGCTGAGCTAGT 896
QY 582 ggcgaagaagctctgtcatcaatccgaacagtcataagcctttagggcccaaggaaga 641
DB 897 AGCTGAAGAGTCTTGGCAGCAATTTAGAACGTGTATTTGATGAGAGCAAAAGAGA 956
QY 642 acttcaaaagtatcacagaatcacaagatgcagaagatlttggcatataaaggactat 701
DB 957 ACTTGAAGGTACAAACAAATTTAGAAAGAACTAAAGAAATTTGGATTAAGAAAGACTAT 1016

```



```

Db 3171 CATGAAAAAACCCCTTGATGACAGCTACAGCAGGAAGGCTTAATGCCAGACATTT 3230
OY 2880 tgaagggaattagagtttcgagaagtcctctctcatcatcgtgcgccagatgttt 2939
Db 3231 GGAAGGAATATGCACATTTGGTGAAGTGTATTAACATATCCACCCGACCGGACATCCC 3290
OY 2940 catctcgttcgcttcctccatcagatctgagcgaggaagacagatcagcttgaggag 2999
Db 3291 AGTGTTCAGGAGACTGAGCCTTGAGGTGAGAGAGGCGCCAGACGCTGCTGTGGTGGCAG 3350
OY 3000 cagcgctgtgagaaagacacttcctcgaacttcgacagagacttataagccccgtgca 3059
Db 3351 CAGTGCCTGTGGGAGAGACAGAGTGCTCAGCTCTGGAGCGGTTCTAGACCCCTTGCC 3410
OY 3060 aggaacagtcgtctgttgatgtgtgtgacaaagaattgaatgtacagtgctccgttc 3119
Db 3411 AGGGAAGTGCCTGTGATGCGCAAGAAATAAAGACATGAATGTTCATGTGCTCGAGAC 3470
OY 3120 ccaatagcaatcgtcttcctaaagacccgtgtcttcacatcagacatgtcgtgagaat 3179
Db 3471 ACACCTGGGCACTGTCCTCCAGAGCCCAATCCTGTTGACTGACGATGCTGAGAACAT 3530
OY 3180 cgcctatgtgacaacagccgtgtgtgctccatagatgagatcaaaagaagccgcaaatgc 3239
Db 3531 TGCCATATGAGAGACACAGCCGGGTGTGTACAGGAGAGATTTGAGGGCAGCAAGA 3590
OY 3240 agcaaatcctcatcttattatgaagtcctccctgagaataatacaacaagaattgagct 3299
Db 3591 GGCCAACTACATGCTTCATGAGTCACTGCTAATAAATATAGCACTAAAGTACGAGA 3650
OY 3300 gaaaggaacagcgtcttcgtgagcgacaaagaagacagctatgtgagaaggtctctc 3359
Db 3651 CAAGAAGACTAGCTCTCTGTGGTGGCCAGAAACACAGCATTCGCTATAGCTGTCCTTGT 3710
OY 3360 ccaaaaccccaaatcttattatgttgatgagccactcagccctcgataatgacagtga 3419
Db 3711 TAGACAGCCTCATATTTGCTTTGGATGAAGCCACGCTGAGCTGTGATACAGAAAGTGA 3770
OY 3420 gaagtggttcagcatgcctctgataaagccagagcggaagacatgcctcgtgtgacg 3479
Db 3771 AAAGGTTGTCCAAAGAGCCCTGACAAAGCCAGAGAGGCCGACCTGATGTGATTGC 3830
OY 3480 tcaagcgctctcgtcagatcagaagcagatctgagtgatgtctcgcagatgagaagat 3539
Db 3831 TCACCGCCTGTCCACATTCACAGATGACAGACTTAATAGTGTGTTCCAAATGGCAGAGT 3890
OY 3540 aaaggaacaagaactcatcaagagctcctcgaagaatcagagacataatcttaagttagt 3599
Db 3891 CAAGGAGCATGGCAGCATCAGCAGCTGTGGACAGAAAGCATTTATTTTCATGTGT 3950
OY 3600 gaatgcacagtcag 3613
Db 3951 CAGTGTCCAGGCTG 3964

```

RESULT 8  
US-08-784-649A-5  
Sequence 5, Application US/08784649A  
Patent No. 5830697

GENERAL INFORMATION:  
APPLICANT: SIKIC, Branimir I.  
APPLICANT: Chen, Gang  
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
TITLE OF INVENTION: CYCLOSPORIN MODULATION  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-784-649A-5

```

Query Match 36.0%; Score 1303.6; DB 2; Length 4264;  
Best Local Similarity 61.9%; Pred. No. 0;  
Matches 2162; Conservative 0; Mismatches 1284; Indels 48; Gaps 4;

```

OY 162 caagtgacccctgcatatgtagtggaatgtagtgccttgatctttagtatacaca 221
Db 477 CAGGTATGCTTATTTATTTACAGTGAATGTGCTGTGGGTCCTGCTGCTTACATTTCA 536
OY 222 gattccttgtagtataactgacagcagacagacagagagatcgcgaagaagtttt 281
Db 537 GGTTCATTTTGGTGGCTGCGCACCTGGAAGCAAAATACAAATTAAGAAACGTTT 596
OY 282 tcatcagtttgagcagacacatcggttgtagtgcacgtgacatgcgttgagagctaa 341
Db 597 TCATGCTTAATATGCGACAGAGATGAGGCTGTGTTGATGTGCACGATGTTGGAGCTTA 656
OY 342 cactcgcatgacagacatgacacaaatcagtgatgtagatgagataagaattgctctgt 401
Db 657 CACCGACATTACAGATGATGTCTCCAAAGATTATGAAGAAATGGTGACAAATTTGAAT 716
OY 402 gttcacaacatgctacttcttcgatltggcctgagcagttggttgtagagggctgga 461
Db 717 GTTCCAGTCAATGCAACATTTTCACTGGGTTATAGTGAATTTACACGTGGTTGGA 776
OY 462 actcaccctagtagactctacacagctcctcctctataatgctcgaagcagcagatgc 521
Db 777 GCTAACCTTTGTGATTTTGGCCATCAGTCTCTGTTCTTGACGTGCACTGCTGTGCGC 836
OY 522 taggatgcatctcatcttgacacagtaagtaagtgccatctccaagctggggctgt 581
Db 837 AAAGATACATCTCTCATTTTCTATTAAGAACCTTAAGCTATGACAAACCTGGAGAGT 896
OY 582 ggcagaagaagctctgcatcnaatccgaacagtcataagccttaaggccagagaaga 641
Db 897 AGCTGAAGAGGCTGTGGCAGCAATTTAGAACCTGATGTGATTGGAGGACAAAGAAAGA 956
OY 642 acttcaaaagtatcacacagatctcaagaatggaagatlttggcgttaaaagagactat 701
Db 957 ACTTGAAGGTACACAAATTTAGAGAGAGCTTAAGAAATTTGGATTAAGAAAGAACTTAT 1016
OY 702 agctcaaaagtgctctctgtgtgtgtgtactctcttataatgaaagacactgactgtc 761
Db 1017 TACAGCCAAATTTCTATAGTGTGCTGCTTCCGCTGATCTATGACATCTTATGTCTGGC 1076
OY 762 ttlttgtagaacctcctcttgatcttcaatgagaaacctgagataccatcggagctgt 821
Db 1077 CTTCGTGTATGAGCACCTTGTGCTCCTCAGGGGA-----ATATCTATTTGACAAAGT 1130

```



Db 3291 AGTGTTCAGGAGCTGAGACCTGGAGGTGAAGAAGGGCCAGACGCTGCTTGTGGGAG 3350  
OY 3000 cagcggtctgaggaaagcaactctgttcaactcttcagagaaacttattgacccctgca 3059  
Db 3351 CAGTGGCTGTGGAGAGACACAGAGTGGTCCAGCTCTGGAGCGGTCTACAGACCCCTTGGC 3410  
OY 3060 aggaacaatgctgttattgattgattgattgattgattgattgattgattgattgattg 3119  
Db 3411 AGGGAAGTGTGCTTGTGATGAGCAAAATTAAGGACTGAATGTTCAGTGGCTCCGAGC 3470  
OY 3120 ccaaatagcaatcgtctcccaagagccgtgtcttcaactcagcagcattgctgagaaat 3179  
Db 3471 ACACCTGGGACCTGTGTCCAGAGACCCATCTGTTTGACTCTACAGATTGCTGAACAT 3530  
OY 3180 cgcctatgtgagcaacagccgtgtgtgagccattagattagattgattgattgattgattg 3239  
Db 3531 TGCCATGTGAGACACACACCGGGGTGTGACAGGAAGATTGTGAGGGCAGCAAGGA 3590  
OY 3240 agcaaatatcatcttatttattgaaaggtctccctgagaaatcaacaagaattgagat 3299  
Db 3591 GGCCAACTACTGCTTCATGAGTCACTGCTAATTAATTAATTAATTAATTAATTAATTAAT 3650  
OY 3300 gaaagagacacagctcttcgtgagccagaaacaaagactagctattgcaaggctctct 3359  
Db 3651 CAAGAGACTAGCTCTCTGTGTGGCCAGAAACAGCATTTGCCATAGCTGCTGCTTGT 3710  
OY 3360 ccaaaaccccaaaatttatttattgattgattgattgattgattgattgattgattgattg 3419  
Db 3711 TAGACAGCTCATATTTGCTTTGATGATGAAGCCACGTCAGCTGATGATGATGATGATG 3770  
OY 3420 gaagtgattgagcagctctgtatgaagccagagcagaggaaggaaggaaggaaggaag 3479  
Db 3771 AAAGTTGTCCAAAGAGCTGTGAGCAAGCCAGAGAGGCGCCACCTGATTTGATTTGC 3830  
OY 3480 tcaacagctctctgcaattcacaacagcagattttagattgattgattgattgattgattgattg 3539  
Db 3831 TCACGGCTGTCCACATCCACATCCACATCCACATCCACATCCACATCCACATCCACATCC 3890  
OY 3540 aagaagacaggaactcacaagctccctgagaaatcgagacataattttaaattgattg 3599  
Db 3891 CAAGAGAGATGAGCAGCATGAGCTGCTGAGCAAGAAAGCATATTTTCAATGAT 3950  
OY 3600 gaatcacagatcag 3613  
Db 3951 CAGTGTCCAAGCTG 3964

RESULT 9  
; Sequence 1, Application US/09120513  
; Patent No. 6025160  
; GENERAL INFORMATION:  
; APPLICANT: Brun, Kimberly  
; APPLICANT: Chenery, Richard  
; APPLICANT: Eliens, Harma  
; APPLICANT: Field, John  
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Smithline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY:  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/120,513  
;; FILING DATE: 22-JUL-1998  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: King, William T  
;; REGISTRATION NUMBER: 30,954  
;; REFERENCE/DOCKET NUMBER: GP50008  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-270-5015  
;; TELEFAX: 610-270-5090  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4233 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: Genomic DNA  
US-09-120-513-1

Query Match 33.5%; Score 1212; DB 3; Length 4233;  
Best Local Similarity 60.4%; Pred. No. 0;  
Matches 2114; Conservative 0; Mismatches 1340; Indels 48; Gaps 5;

Db 162 caggttgaacctgtattatgtgaaataggtgtgtctgcttgaatttggttacatata 221  
OY 162 caggttgaacctgtattatgtgaaataggtgtgtctgcttgaatttggttacatata 221  
Db 355 CATGTACGCTCTACTATTACACGCGCATTTGCTCCGGTGTCTATCAGCTTGCCTACATCA 414  
OY 222 gatttccttgattatataatcagcagcagcagcagcagcagcagcagcagcagcagcagc 281  
Db 415 GGTTCACCTTGTGTGCTGCTGAGCTGGAGACAAATACAAAGATTAAGGCAAGATTTT 474  
OY 282 tcaatcagatttggcacaagacacagcagcagcagcagcagcagcagcagcagcagcagc 341  
Db 475 CCATGCCATCATGATCATCAGAGATATAGCTGTGTTAGCTGATGACGCTGGGAGCTCA 534  
OY 342 cactgcagatgaca--gacatgacaaatcagtgatgattgagatgagatgattgctct 398  
Db 535 CACCGGCTCAGACAGATGAGCTGCCAAATTAATGAGAAATTTGGACAACTTGGAAAT 594  
OY 399 gttgttcaaaacatgctcattcttctgattgctgctgctgctgctgctgctgctgctgctg 458  
Db 595 GTTCTTTCAGTCCATTAACGACATTTTACGCCGTTTATTAATGAGATTTAATGATGCTTG 654  
OY 459 gaaactcaacctagtgactctacacagctcctccttataatgagctcagcgagcagcagc 518  
Db 655 GAAGCTAACCTTGTATTTTGGCCCTCAGCCCTCTTATTGGGTGTCTATGCGCATGTG 714  
OY 519 ttctagatgctcattctatcattgacagtaaggaattgaattgcttattcctcaagctggggc 578  
Db 715 GGCAGAGTACTGACTTCAATTAATTAAGAACTCCAGGCTTATGCGAAAGCTGAGAC 774  
OY 579 tgtgacagaaagatcttctatcaatccgaacagtcagcttgaagcccgagagaga 638  
Db 775 AGTTGCCGGAAGATCTTTCAGCCATTCAGAACTGTGATTTGGTTGGAGCAAAAAAG 834  
OY 639 aagaactcaaaaglataccagaatcacaagatgcaaggaatttggcataaaaagagc 698  
Db 835 GGAAGCTTGAAGGTACATTAATAAATTTAGAAAGCTTAAGAAGTGGCATTAAGAAAGC 894  
OY 699 tatagcttcaaaagtgctcttggt 758  
Db 895 CATCAGCGCCACATTTCCATGAGTATTCCTACCTGTTGGCTATGCTTATGCTTATGCTACT 954  
OY 759 tgcctttgtgtgagacactccttgaattcttaatgagaaactgagatatcaactcggagc 818  
Db 955 GGCATTTCTGTATGAGACCTCCTTGTGTCT-----CTCAATGATTAATTTATTTAGACA 1008  
OY 819 tgtctctgtgtcttcttgaattgataatcagcagttatgcaattgagcagcagctccc 878  
Db 1009 AGTGTTCAGGCTTCTCTTATTTATTTATTTGAGGACTTTCAGTATTTGACATTTAGCCCC 1068



879 tcacttligaaccctcgcaataagcccgagagctgccttccatattcccaagttatga 938  
1069 AACATAGAAAGCCTTTGCAAAATGCAAGAGGGGAGCCTATGAAATCTTCAATTAATTGA 1128  
939 taagaaaccagtatagataaacttccacagctgatatataaacctgaatccatagaag 998  
1129 TAATGAGCCAAAGCATGTGACAGCTTCTCAACCAAGGAGCAACACAGACGATTAATGGG 1188  
999 aactgtggaatttaaaatgcttcttccaaatcatcaagaacccatctcaagaattct 1058  
1189 AAATTTGGAATTTAAATATGTTACTTCAACTCCATCAAGAGTGAATTAAGATCTT 1248  
1059 gaaaggtctgaatcagaatlaaagctcgtgagagacagtcgccttgcgttcaatgg 1118  
1249 GAAGGGCTTAACCTGAGAGTGAAGAGCGGAGAGCGTACCGCTGGTTGGCAACACTGG 1308  
1119 cagtcggaagagtagcagtagcagcttctcagaaggtatatacgatccggaatgagct 1178  
1309 CTGTGGGAAAAGCACAACCTGTCAGCTGTCAGAGGCTTACGACCCCATAGAGGGCGA 1368  
1179 tatcatggtgagatgagatatacatcagaagctttaaagtgcgcatatcagaacat 1238  
1369 GGTCAATATGACGAGACAGAGACATCAAGACCATCAATGTAAGTATCTGCGGGAATCAT 1428  
1239 tggagtgatgaatcagaagagctgttctgttcgagagacacatcagtaacaatcaagta 1298  
1429 TGGGGTGTGATGAGAGAACCCGTCGTTTGGCACACAGATTGCCAAAACATTCGCTA 1488  
1299 tggagagatgagatgagatcagaagagatgagagagcagcaagaggaagaatgagta 1358  
1489 TGGCGGAGAAAACCTCAACATGATGAGATAGAGAAAAGCTGTCAAGAGCAATATGCTTA 1548  
1359 tgaattatcatgagagttccctaataattatatacttgtaggggaaaaagagctca 1418  
1549 TGACTTATCTATGAAACTGCCCCACAATTTAACAACCTGTTGTTGAGAGAGGGGCGCA 1608  
1419 aatgagtgagagagagagagagagatcgcaatgctgcgttgccttagtltcgaaacccaa 1478  
1609 GCTGAGTGGGGGAGAGAAACAGAGATGCGCATTCGCCGGGCCCTGGTCCCAACCCCAA 1658  
1479 gattctgaatttagatgagctacgtctgcctcgtgattcagaagaagcaagtaagctgttca 1538  
1669 GATCCTTTGTTGGATGAGGCGACGTCAGCCTGTGACACAAAGAGCGAAGCGTGTGTCA 1728  
1539 agctgcactgag 1598  
1729 GGCGGCTCTGATTAAGCTTAAGAGAGGCGGAGACCATTTGTGATAGCTCAACGCTTGTG 1788  
1599 tactattcgaagtgagatltgattgtgacccaaaggaatggaatgctgcgagagaaag 1658  
1789 TACAGTGGCGCAATGCTGACATGCTGCTGTTGATGTTGATGTTGATGTTGAGAGCAAG 1848  
1659 agcacatgctgaactaaatgagcaaaagagagagagagagagagagagagagagagagag 1718  
1849 AAATCATGAAGAGCTCAAGAAAGAGAGAGGAGGCAATTTACTTCAAACTTGTCAATGACACAGAC 1908  
1719 taataaaaagctgagta--agagtgagatcagatcatctcactaagaagaagacaa 1777  
1909 TAGAGGAATATGAATTAACCAAGAAATATGCTTATGATATGCCAAAGTGAACACTGCTGTC 1968  
1778 actcaactcctcgtcactcgtgaaagagatcaagatcagaacttcatgcaagagctgag 1837  
1969 CTCTAGTGTGACTTCAGAAAATCAAAATCTCTTAATAGAGATCATCAATTCGAGAGAG 2028  
1838 aatcaaccaatctaa-----agagataag 1862  
2029 TATCCACAGAGAGACAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2088  
1863 tctctcgaagctctctatataaaatttaaaagttaaagctgaagctgagcttctgt 1922  
2089 TGTGCTATGTTGCTTTTGGCAGATCTTAAGCTAAATATTAATGATGATGAGCCCTATTT 2148

1923 ggttcgtggggaacttgctctgttcttaaatgaaactgttccatccagtaatttccatcat 1982  
2149 AGTTGGGGTACTTGTGCTGTTATTAATGGTCAATACAAACAGTGTGGCCATTA 2208  
1983 ctctgcaaaatataaccatgctt--ggaataatgataaaacacatlaaagatga 2039  
2209 GTTTTCAAGATGTAGGGGTTTTTCAAGAGACGACGACCATGAACCAACAGGAA 2268  
2040 tgcagaattatccatagatattcgatattgggtgtattgttattgttgcgaattatt 2099  
2269 TTGTAACTTGTTCCTTCCTTCTTCTGTGATGGAAATGATTTCTTTGTTAGTACTT 2328  
2100 catcgagagatatttcaagcagagcagaggaattttaaagatagatlaagaaact 2159  
2329 CTTTCAAGGCTTACATTTGGCAAGCTGAGAGATCTCACCAGAGCTCGGATACAT 2388  
2160 ggcctcaaaagcatttatacagatatctgcgttgttgaatgaagaagaagaagcagc 2219  
2389 GGTCTTAATTCATGCTGCGACAGATATAGCTGGTTGATGACCATTAATAAACACAC 2448  
2220 aggaagcttgacaacaataltagccatagatatagcaacaattcaagagcaaggttc 2279  
2449 TGGCTCGCTACTACACAGGCTCGCTAGTGAAGCTTCAATGTTAAAGGGGCTATGGGCTC 2308  
2280 cagagatggcgttcttacaacaataatgacaactaagagacattcagtaattccatt 2339  
2509 CAGGCTTGTGTAATGTAACCAAGATGTAACCAACCTTGGCAGAGATTTATCTTATCTT 2368  
2340 tatalatgagtgagagatgacatccatcctgattctgagatgttccagtaactgcgtgac 2399  
2569 AGTGTATGGCTGGCAGCTTACCTTTACTTGTGATTAATATACACTATTTGCTTGGG 2628  
2400 aggaatgattgaaacgcagcagcaatgactgatttgcacaagaagatagaagaactaa 2459  
2629 TGGAAATTTGAAATGAAATGTAACCTGTGTGTCAGAGCCTTGAAGAGCAAGAAAGGCTAGA 2688  
2460 gcatcgtcggaaagatagcaatgaaacttggagagacatacgtatataagtgatcaat 2519  
2689 GATCTTGGGAGATTCGCTACAGAGCAATTTGAAACTTCGCACTGTGTGCTTTGAC 2748  
2520 aagggaaaagccttcgagcaaatgatatgaagagatgcttcaagactcaacacagaatac 2579  
2749 TAGGAGAGAGAAATTTGAATATGATGATGCCACAGCTTGACATACCAAAATGTC 2808  
2580 ctgagaagaagacagatattatggaagctgttatgcatcagccatccatataatt 2639  
2809 TTTGAAGAAACACACGCTTTGGATCCTTGCCTTACCCAGGCGCATGATTTATTT 2688  
2640 tgcctatgcaagcaggttgcatttggagcctattttaaactgaagctgagcaatgacccc 2699  
2869 TTCTATGCTGCTGTTTCCGTTCCGTTCCGTTCTACTTGTGTCAGAGAACTCATGACGTT 2928  
2700 agagggcatgtcatagtttactgcaatgcatatgagagatagagatagagagagagagag 2759  
2929 TGAATAATGTAATGTTGATTTTCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2988  
2760 gctcgttctgctcctgatatccaaagccaatcggggctgagatcgttgcctt 2819  
2989 CAGTTCAATTCCTCTGCTACGCGAAGGCCAAAGTCTCACATCTCCATCATCAGAT 3048  
2820 gttggaagaagaacaataatagacagcgcagatcaagaagaagaagaagaagaagaacatg 2879  
3049 CATTGAGAAATCCCGAGATTTGACAGCTACAGCAGAGAGGCTTGAAGCCTTAATGTTT 3108  
2880 tgaaggaatttagatlltcgagaagcttcttcttcatccatcgtgcgcaagatgltt 2939  
3109 AGAAGAAATGTAATTAATGAGTATGTTCAATATCCACCGAGCAACCAATATCC 3168  
2940 catctcgttgcgttataccatcagatgagcagagagaagaagaagaagaagaagaaga 2999  
3169 AGTGTTCAGGAGCTGAGCTTGCAGGTGAAGAGGGGCAAGGCTTCCCTGTGTGGAG 3228  
3000 cagcgctgttggaagaagactcgttcaactctgcagagacttataagaccctgtga 3059



```

Db 3229 CAGTGGCTGGCGGAGAGTACACTGGTCCAGCTGCTCGAGCGCTTCTACAAACCCCATGGC 3288
OY 3060 aggaacaagctcgtctttagatggtgtgatacgaaagaatltgatatgacgttgcctcgtc 3119
Db 3289 TGGAAACAGTGTCTTCTAGATGGCAAAAGAATTAATAACACATCAACAGTCAGTGGCTCGCGC 3348
OY 3120 ccaaatgacatcgtctccaaagagcgtgtgtcttcaactcgcacgtcagcatctgtgaaacat 3179
Db 3349 CCACCTGGCATTGTGTGCCAGAGCCCAATCCTGTTTGAATGACACATCAACCGAAGACAT 3408
OY 3180 cgccctatgctgacaaacagccgtgtgtgtgcctatagatgatacaagaagccgcaaatgac 3239
Db 3409 CGCCTACGGAGACACACGCGTGTGCTGTCTATGAGGAGATCGTAGAGGCCGCCGCAAGGA 3468
OY 3240 agcaaatatccattctttatgtgaaggtcctctgagaataacaacacacagttgact 3289
Db 3469 GGCCCAACATCCACACAGTTCATGCACCTGACCTGAGAAATTAACAACACAGTGGGAGA 3328
OY 3300 gaaagagcacaagctcttcgacgacgaagaacaaagactagctatgtcaagagctctct 3359
Db 3529 CAAGAGGACCTCAGCTGTGCGGCGCGGAGAGACAGCGCATCGGCATCGCGCCGCCCTCGT 3588
OY 3360 ccaaaaccccaaatltaattgtltgatagagccacttcagccctcgataatgacatgta 3419
Db 3589 CAGACACCTTCACATCTTACTTCTGTGATGAACGACATCAGCTGTGATAGGAGAGTGA 3648
OY 3420 gaagatggttcagcatcagcctgtataaagccagagcggaagagacatccctagtggtcac 3479
Db 3649 AAGAGTGTCTCAAGAGAGCGGTGAGCAAAACCCAGGAGAGGCCACCTCATTTGATGATGC 3708
OY 3480 tcaaacgctctctgcacatttaagaacgcagaattltgatagtgtcttcgacaaatggaagat 3539
Db 3709 GCACGCGCTGTCCACCATCCATCAGAAACGACAGACTTGATCGTGGATTCGAAGAGCGCAGGT 3768
OY 3540 aaagagcaaggaacatcatalcaagagctcctctgagaagaatcgagacatatatttaagta 3599
Db 3769 CAAGGACACAGGACCCACACGACGCTGTGGCCCGAGAAAGGCATCTATTTCTCGATGCT 3828
OY 3600 gaatgcacagtcagtcagtga 3621
Db 3829 TCAGGCTGAGCAAGGCGTCA 3850

RESULT 10
US-09-450-105-1
: Sequence 1, Application US/09450105
: Patent No. 6169166
: GENERAL INFORMATION:
: APPLICANT: Kimberly Anne Brun
: APPLICANT: Richard James Chenery
: APPLICANT: Harma Eilens
: APPLICANT: John Anthony Feld
: APPLICANT: Lin Yue
: TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
: TITLE OF INVENTION: ENCODING RAT MDRLB2 AND SCREENING METHODS THEREOF
: FILE REFERENCE: GP-5008-D1
: CURRENT APPLICATION NUMBER: US/09/450,105
: EARLIER FILING DATE: 1999-11-29
: EARLIER APPLICATION NUMBER: 09/120,513
: NUMBER OF SEQ. ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 4233
: TYPE: DNA
: ORGANISM: HOMO SAPIENS
: US-09-450-105-1

```

Query Match	33.5%	Score 1212;	DB 4;	Length 4233;
Best Local Similarity	60.4%;	Pred. No. 0;		
Matches 214; Conservative	0;	Mismatches 1340;	Indels 48;	Gaps 5.

QY	162	cagtgtagccctgtatatagttggaaagvtgtctgccttgattttgtttacataca	221
Db	355	catgtagccctactatataccacggcatttggtcgggtgtgtcatctgtgctctacatacca	414
QY	222	gatttcctgtgagttataatctgcagacagcagaccgaagagatttcgaaacagttttt	281
Db	415	ggttccactttgtgcctgcgcagctgggagacaacataccaaatattagcgaaagttttt	474
QY	282	tcaattcaattttgtgcagacagatccgctgtttgtatagctgtacatccgtggaacttaa	341
Db	415	ccatgcatcatatgatatacgaagataagctgttcttgacgtgaatgaagcctggggagctcaa	534
QY	342	caatgcgatgaca--gacattgacaaatcagtgatgbatgtgagataagattgctct	398
Db	535	caccggctccagatgacgtctcccaaatattagacgaattgtgtgacaaacttgat	594
QY	339	gttttttcaaacattgtcactatttttgatattgcccgtgcagtttggtttgtgaaggctg	458
Db	595	gttcttcttgcataccatacagacattttcagccggtcttataatagatttataagctgtgtg	654
QY	459	gaaatccccccttagtgactatccacagctctccctcttataatgatctgcagcgacagcaty	518
Db	655	gaagctaaacctgttaatttggccgttcagccctcttatttggttgtgtcatctcgcaatgy	714
QY	519	tctctagatgtgtcatctcatatgacagtgaaagaaatlaagtgctatltccaaagctgggac	578
Db	715	ggcaaaagttactgactcttacttaataagaatccagatccaggtcttagcgaaagctggagc	774
QY	579	tgtgtgcagaagaagttcttgttcatcaatccgaacagttcaatagccttagtgcccagggaa	638
Db	775	agttccgaagaagctcttagcagacatacgaactgtgattgctgttggagcacaagaagaa	834
QY	639	agaacttcaaaaggtatcacagacgaactcaaaagctgcaaaagatttggagataaaagac	698
Db	835	ggaactgtgaagaagtagacataaaattagaagaagcctuaaagaagtgtgacataagaagaac	894
QY	699	tatagcttcaaaaagtgctctctgtgtgtgtgtactcttattatgaaatggaacctatgact	758
Db	895	catcacggccaaactttccataggtatgtgcacactgtgtgtctatgctgtctatgact	954
QY	759	tgtcttttgtgtagaacctcctctgattcttcaatgtagaactgtgatacatcogggac	818
Db	955	ggcattctcgtatagagccctcctgtgcct-----ctcaaatgaaatattcatatttgaca	1008
QY	819	tgtcttgcgttttcttttagtgaatcccatagcagttatgtgattggagcgagcgctcc	878
Db	1009	agttgttaacgctctctctctcatatttatttggggacttcaaglatgtgacatttggcccc	1066
QY	879	tcaacttgaaacctctgcgaatagccgaagagagctgccttcatatttccaggtattatga	938
Db	1066	aaactatgaagccttttgcgaatgcaagagggcgagcctatgaaatttccaagataattga	1122
QY	939	taagaacaaccgtagatagataaacttttccacagctgtagataaaccctgtaacatagaag	998
Db	1122	taatgagccaagcattgacagcgttcttcaaccaaagggaacaaacagacagatataatg	1188
QY	999	aactgtggaatttaaaaatgtttcttccaattatccatcaagacacatcatcaagaattct	1055
Db	1189	aaatttgaattttaaaaaatgttttacttccatactccctacagaaatgaaagttlaagattctt	1246
QY	1059	gaaaggtctgaaatctcgaatttaagctctggagagacagctgcctgtgtgcgtctcaatgy	1118
Db	1249	gaaaggcttcaaccgtgaagggtgaagagcgggcagacggttagccctgtgttgcgaacggtg	1306
QY	1119	cagttggagaagtagacgtgacagctcttcgcagaggttatagataccggatgattgct	1176
Db	1309	ctgttgaggaaagacaacactgtccagctgcgtcgcaagagcctctacgaccatagaggcgga	1366
QY	1179	tatcatgtgtgagtagaagatgacatccagagctttaaagtgtgcgcattatccgagacacat	1233
Db	1369	ggtccagatctgacgagcagcagacatcagagacatcaatcaatgtgaagttatcttcggyaaatct	1426

QY	1239	tgnagtggttagtcaagagccctgttttctgttgggaccacatcagtaataacatcaagta	1298
Db	1429	tgggtgtgtgtagtcagagaaacccgttcgttttgcgccacgattgcggaaaaatttcgta	1488
QY	1299	tggagagatgtatgtgacatgcatgaagagatgtggagagacagcaaggaagcaaatcgta	1358
Db	1489	tggccgagaaaaacgltcaaccaatgta-gtagatagaagaagctgttcaaggaaagccaatgccta	1548
QY	1359	tgatttcatctgtagaggtttccctaataattatacatctgttaggggaaaaagagatca	1418
Db	1549	tgaattcatcatgaagaactgccccaacaatttaacaccctgtgtgtagagagggcgca	1608
QY	1419	aatgtgtgtagggcagaaacagaagatcgcgaattgtctgtgctttagtcgaaacccca	1478
Db	1609	gtctgtgtgggagacaaacagaagatcgcgaattgtcccggtccctgttccgcaacccca	1668
QY	1479	gatttcgattttagatgtagagctgacgtctccctggatttcagaagaagtcagcttca	1538
Db	1669	gattccttctgttga-gaagagccaagctcagcctctgtgacacagaagaagccgtgttca	1728
QY	1539	agctcacatgtagaagcgagcaagaagtgctgacataactcgttgaagcacaacccgatttc	1598
Db	1729	ggccgctctcgataagagctcagagaagggccggaccacattgtgatagcacacgctgttc	1788
QY	1599	tactattcgaaagtgcgaatttgatgtgacccctaagaagatgaaatgctgcggagaaga	1658
Db	1789	tacagtgcgcaatgctgcgaactcatctctgttttga-tgtgtgtgattttgttgagcaag	1848
QY	1659	agcacatgtctgaactatgcaaaacagagctcataltttaactctgttgatgcacagga	1718
Db	1849	aaatcatgaaagctcatgaaagaagaagggcattcttcctaactgttcatgacacagac	1908
QY	1719	tattaaaaagctgtaagaa--cagatgtagatcgaatgacataattctactgaaagaagacca	1777
Db	1909	ttagagaaatgtaaatgtgaaccaggaaataatgatttagatcccaagtgcactgtgc	1968
QY	1778	actcaactcctctgcactctgtgtgaagagcatcaatgcagatctcatctgnaaagctgag	1837
Db	1969	ctctgagttgtaactcggaaaaatacctaactctcctttaagaagatcaattgcgcagaag	2028
QY	1838	aatcacaccacataa-----agagataag	1862
Db	2029	tatcacacagaagaacagaccagagagagaaacttaagtccaagaagatgtgtgatagaa	2088
QY	1863	tctcttgagagtcctcctataaataatttaagaatttaaaacagcctgaaatggtccttgt	1922
Db	2089	tgtgcctatggttcccttcttggaagctccctaagaagctcaatatagtagagccctattt	2148
QY	1923	gattctgtaggaacttgcctctgttctaaatggaactgttcacatcttccatcat	1982
Db	2149	agttgtgggtgtactctgtctgttataaattgggtgcatacaacagttgttgcatagt	2208
QY	1983	ctttgaaaaatatbaaccgttt---ggaataatgataaaaaccacataaagctga	2039
Db	2209	gtttcaaaagatgttaggggttttlttcaaaagcagcagaccatgaaaccaaacagga	2268
QY	2040	tgcagaatttatctcatgatactgcatacttgggtgatttatctgttcgttgcagtatt	2099
Db	2269	tgttaactgtttccctctcttcttcgtgcataggagatgattcttctgtttagtaact	2328
QY	2100	catgcagggatatttbaogcgacagcaggggaaatttaagatgagatbaacaact	2159
Db	2329	cttcaaggttcaacatttgcgaaagtgtgagatcctcaacaagcactcgcatacat	2388
QY	2160	ggccttcaagggcatgttatacatgaattgtcgtgtttgattgaaagaagaaaacagac	2219
Db	2389	ggtcttcaaatcgaatcgtcgaaagatataaagctgtgttgatgacataaaaaacacac	2448
QY	2220	agtagagtttaacaacaatatctagcatagataatagcaacaattcaaggaacaagtttc	2279
Db	2449	tgtctgcgtacacacagcgtcgtagtgcagcttccaatggttaaaagggctatgggttc	2508
QY	2280	cagagattggtgttcaacaaaatgaaataacatgaacatgagacttcaattcalcttcc	2339

Db	2509	caagcttcgctgtgttaaccagaatgttagcaaaccttgcgcacgaattatctatccctt	2568
Qy	2340	tatatcgtacgtggagatgtagcattccgtatcttcgtgatctgcaccagtaacttgcgttagc	2399
Db	2569	agctatgctgcgcagcttacacacttctacttctgttagtaattatacactcattgctctggg	2628
Qy	2400	aggaatgattgaaacccgcagcaatgtacgtgatttgcacaagaatagaagaagacttaa	2459
Db	2629	tgtatatttggaaatgaaaactgttctctgtcgttgcataaagccttgaagacagaagaagctaga	2688
Qy	2460	gcatagcgtgaaagatgcaactgaagccttctggagaatactagactatagtgatcataac	2519
Db	2689	gactccggtggaagatcgtctacagaagaacttgaaaaacttcgcgactgttctctttggac	2748
Qy	2520	aagggaagaagccttcgcagcaaaatgtatgataagagaagtctatagaactcacagaataac	2579
Db	2749	tagggagacagaatgttggaaactatgtatgcccagagcttgcagatatacagaaatgc	2808
Qy	2580	ctcgaagaagacacagattatgtgaaagcgtgttaagtacatgcagccatgctttatattt	2639
Db	2809	tttgaagaagaacacgctcttttggatctacacttcgccttcaccccgagccatgattattt	2868
Qy	2640	tgctatgcaagcagaggtttcagatttggagccttattaactaaagtgtgaacgaatgagccc	2699
Db	2869	tctcctatgcgtcgtgttccgcgttccgtctcgtcacttgcattgtgtgcagagaaaccatgagct	2928
Qy	2700	agagagcattgtcctaaattcttcttcacaaatgcatactggagctatgagccatgcgaaaaac	2759
Db	2929	tggaaaagtatagtgtgtattcttccgtctgttgccttgcgtgcagtgagcaggaagaaac	2988
Qy	2760	gctcgttttgcctcgtgaataatccaaagccaaatccggggtcgtgcacatcgtttgctt	2819
Db	2989	cagttcatctgcctctgaactacgcggaagccaaagtctcaagatcccatccatccagat	3048
Qy	2820	gttggaaaaagaaacaaatatagacagccgcgcagctcaagaaggagaaagacagacacatg	2879
Db	3049	cattgagagaaaatcccgagatgtcacgcctacacagacgcgggcttggaaagcctaattgct	3108
Qy	2880	tgaagaggaattagaagttctcgaaagctctcttctctctatccatgtcgcgccagatgtctt	2939
Db	3109	agaagagaaatgtgaaatttaattgaatgtcatgttcaactatccaccgcgacccaatcc	3168
Qy	2940	cactccgcgtgcttatccctcaglatgtgacgaggaagaagacagtagcatttgtfvgggag	2999
Db	3169	agtgttcaggagactgagcttctgaggtggaagagggcacaagctctgcctgtfvgggcag	3228
Qy	3000	cagcgcgtctgvggaaagcacctctctgttaactctgcgcgaagactttatgccccggtgca	3059
Db	3229	cagtgctgcgtgggaagagatacagtgtgtccagctgtgcgcgcgcctctcaaaccccagtcg	3288
Qy	3060	aggaacaagtgtctgttgaatgtgtgagatgcaagaagattgaaatgtacagtgctccgttc	3119
Db	3289	tggaaacagtgttctctagaatgcaagaagaataaacaacactccaagctccaagtgctccgcgc	3348
Qy	3120	ccaaatagcaaatctgtccctcaagaagcctgtgtccttcaactcagctgcagcatgtgtagaacat	3179
Db	3349	ccaaacttggcattgtgtgtcccaagagcccatcctgtttgactgcagatataccagaatacat	3408
Qy	3180	cgctcctagtgtacaacagccggtgtgtgtgcattagtatgagatccaagaagccgcgcaaatgc	3239
Db	3409	cgctcctacggaacaaacagccggtgtgtgtcgtgtctcatgtagagagatcgtgtgagggccgcgaagga	3468
Qy	3240	agcaaaatctcaattcctttatgtagaagttccctctggaagaatcaacaacacagatgtgagct	3299
Db	3469	ggcccaacatccacagttcatcagctcactcgtcgttggaataatacacacagagfvtgggga	3528
Qy	3300	ggaagaagcacagcttctctgcgcgcacagaacaaagaactagctatgtcagaaggtcctctc	3359
Db	3529	caaaagagactcagctcgtcgtgcgcggcgaagaaagcagcgcacatgcacatgcgcgcgcctcgt	3588
Qy	3360	ccaaaaacccaatatattatgtgtgagtagccactcagccctgcgataatgacagtga	3419

Db	3589	cagcgcagccrcaacatcttacttcttcttgatgaagcgacatcagctctcgatagcgagatga	3648
QY	3420	gaagcgcggttcacagatcagtccttgataaagccagacggaagacatgcctagtggtcac	3479
Db	3649	aagcgcgtctccaggaagcgcgttgacaaagccagcggaagcgcgcacatctcatgtgatacgc	3708
QY	3480	tcacagcgcctctctgcaatttcagagacgcagatttgatagtggtctcgacaaaggaaaat	3539
Db	3709	gcacgcgcctcttccacacatccacagaacgcagacttgacgctggtgatctcgaagacgcagct	3768
QY	3540	aaagcgaacaaaggacatcatataaagcctcttgagaaatcgcagacatataattaaagttgt	3599
Db	3769	caagcagcagcgcacccacacagcagctgtctgcgcagaaaggcatcatattcttcgatgt	3828
QY	3600	gaatgcgcagatcagtgacgtga	3621
Db	3829	tcagcgcgcgagcgaagcgcgtcca	3850

RESULT 11  
US-08-461-823-1

```

1 Sequence 1 Application US/08461823
2 Patent No. 5593840
3
4 GENERAL INFORMATION:
5 APPLICANT: Bhatnagar, Satish K.
6 APPLICANT: George Jr., Albert L.
7 APPLICANT: Nazarenko, Irina
8 TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
9 NUMBER OF SEQUENCES: 27
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: OncorPharm, Inc.
12 STREET: 200 Perry Parkway
13 CITY: Gaithersburg
14 STATE: Maryland
15 COUNTRY: USA
16 ZIP: 20877
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.30
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/461,823
25 FILING DATE: 05-JUN-1995
26 CLASSIFICATION: 435
27
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: US 08/168,621
30 FILING DATE: 16-DEC-1993
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 08/010,433
33 FILING DATE: 27-JAN-1993
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Karta, Glenn E.
36 REGISTRATION NUMBER: 30,649
37 REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: 301 208-6997
40 TELEFAX: 301 208-6997
41
42 INFORMATION FOR SEQ ID NO: 1:
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 2726 base pairs
45 TYPE: nucleic acid
46 STRANDEDNESS: double
47 TOPOLOGY: linear
48 MOLECULE TYPE: DNA (genomic)
49
50 HYPOTHETICAL: NO
51 ANTI-SENSE: NO
52
53 US-08-461-823-1
54
55 Query Match 23.6%; Score 856; DB 1; Length 2726;
56 Best Local Similarity 61.7%; Fred. No. 3,2e-250;
57 Matches 1438; Conservative 0; Mismatches 850; Indels 42; Gaps 3

```

```

Qy 2364 cctgattctgagatattgctccagtgacttgcgctgacaggaatgatgaaaccgagcaat 2423
Db 1081 GTTACTCTTGTAGCAATATGATCCATTCATATAGCAGAGTGTGTAAGTAAGAAATGTT 1140
Qy 2424 gactgattgccaacaagaataagaagacttaagcagtcgtaagaagtaagtaactga 2483
Db 1141 GTCTGGACAAACACTGTAATAATAGAAAGAACTAGAAAGTGTGGGAAATCCCTACTGTA 1200
Qy 2484 agctctggaaatacatcgtactaagtgatgatttaacaagaagaaagcccttcgagcaat 2543
Db 1201 AGCAATAGAAATACTTCGCAACCGTGTGTTTCTTGACACAGACAGACAGAAAGTTGAACATAT 1260
Qy 2544 gtaabgaagatgcttcacagctcaacaagaataacctcgaaagaagacagattatttg 2603
Db 1261 GTATGCTCAAGAGTTTGGAGGTACCATATACAGAAACTCTTTGAGAAAGCACACACTTTTGG 1320
Qy 2604 aagctgtatgacttaagcctatgcttattatatttgcctatgacgcaaggtttcgatt 2663
Db 1321 AATTACATTTTCTTCCACCCAGCAATGATGATTTTCTATGCTGAGATGTTCCGGTT 1380
Qy 2664 tggagcctatttaattcaagctgacgaatgacccagagaggaagctgttcatagttttac 2723
Db 1381 TGGAGCTACTTGTGGCACAATAAATCAATGAGCTTTGAGAGTGTCTGTAGATTTC 1440
Qy 2724 tgcattgcatatgagactatggccatcggaaaaaacgctcgttttgctccctgaatttc 2783
Db 1441 AGCTGTTGCTTTGGTCCATGCGCCGCGGCAAGTCAGTTTCATTGCTCTGCTGACTATGC 1500
Qy 2784 caaagccaatcgaggctgagctcgtttgcttgcctgtgttgaaagaagaacaatataga 2843
Db 1501 CAAAGCCCAATATACACACCCCAATCATCATCATATGATCATGAAACCCCTTTGATTGA 1560
Qy 2844 cagccgagctcaagaagaaggaagacacatgtgcaagggagattagatttcgaga 2903
Db 1561 CAGTACAGACAGCAAGGAGCTTAATGCCGAACACTTGGAAAGAAATTCACATTTGGTGA 1620
Qy 2904 agctccttctctatcatcagtgctgcagaaagtgttccatccctcgtgttccatccctcag 2963
Db 1621 AGTTGATTCATCACTATCCACCCGACGACATCCAGTCTTCAGAGCATGAGACCTCGA 1680
Qy 2964 tattgagcgggaagaagcagtagatttftggggaagagagcgtgttgaaagaacacttc 3023
Db 1681 GGTGAAGAGAGGCCAGACGCTGCTGTGGGAGCAAGTGGCTGTGGAGAGACACAGT 1740
Qy 3024 tgttcaactctgcagagacttctaacacccgtgcaggaagcaagtgctgttga 3083
Db 1741 GGTCCACTCTCTGGAGGGGTTTACAGACCCCTTGGAGGAAATGCTGTGATGAGCA 1800
Qy 3084 ggaatgcaaaagaattgaatgtaagtgctccgttcccaaatagcaatcgltccatcaaga 3143
Db 1801 AGAAATTAACGCACTGATGATGTTTCAGTGGCTCCGACACACCTCGGCAATCGTCCAGGA 1860
Qy 3144 ggcctgtcttctcaactgcagcattgtgtgaaacatcgctatgtgtgaaacacagcgtgt 3203
Db 1861 GCCCATCTCTTTGACGTGACAGCACTTGTGAGAACATTTGCTATGGAACAACACCGGGGT 1920
Qy 3204 ggtgcaattagatgagatcaagaagacccgaatgcagcaaatlccattttattga 3263
Db 1921 GGTGTACAGAGAAAGATCTGTGAGGAGCAAGAGAGGCCCAACATATGCTTTCATCGA 1980
Qy 3264 aggtctccctgagaatacaacacacagaatgtgaactgaaggaagcagcttcttgccg 3323
Db 1981 GTCACTGCTTAATTAATATAGCACTAAGTAGAGACAAGAACAGACGCTCTGGGG 2040
Qy 3324 ccagaacaagaactagctattgcaagggctcttcccaaaaccccaaatlttattgtt 3383
Db 2041 CCAAAACCAACGCAATTCATGCTCGTGCCTTTGTAAGACAGCCTCATATTTGCTTTT 2100
Qy 3384 ggaatgaagcaactcagccctcgaaataagatgagagaaggtgttgacgcttgca 3443
Db 2101 GGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAGTTTCCAAAGAGCCCTGGA 2160
Qy 3444 taaagccaggaagggagagacatgctctagtggtcaatcaagcgtctctgcaattcga 3503

```

```

Db 2161 CAAAGCCAGAAAGGCGCAGCTGATGATGCTCACCGCTGTCCACCATCCAGAA 2220
Qy 3504 cgcagattgatagtgcttcgacaatggaagaataaagaagaagcaactatcaaga 3563
Db 2221 TGCAGACTTAATAGTGTGTGTTTTCAGAAATGCGACAGATGCAAGACATGCGCATCAGCA 2280
Qy 3564 gctcctgagaatcgcagacatatatttaagttagtgatgacagtcag 3613
Db 2281 GCTGCTGCGACAGAAAGGCATCTATTTTTCATGATGCTGATGCTGCTG 2330

RESULT 12
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atpd of
; TITLE OF INVENTION: Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
; US-08-996-545-1

Query Match 13.7%; Score 496.8; DB 2; Length 4002;
Best Local Similarity 48.2%; Pred. No. 9.7e-141;
Matches 1705; Conservative 0; Mismatches 1727; Indels 102; Gaps 7;
Qy 172 ctgataatgttgaaatagtggtgtgcttgatttgggttaacatacagatttccctg 231
Db 457 CTGACTTCGATACCTCGGATCGCGAGATTGTCACTGCTATGTAGTACTGTTGGC 516
Qy 232 tggattataactgacagacagacacacgaagagagtttgaaagagtttttcacagtc 291
Db 517 TTCACTATACCGGAGAACACGCGACGACAGATCCGCGAGTATTCCTTGAGTATTC 576

```

OY 292 ttggcagcagacatcgctggtttgtagctgtgacatcggtggaacttaacactgcatg 351  
DB 577 ctggccgccaamaactggtatatttttaacgtcggtggtggggaactgacacccctata 636  
OY 352 acagacatgacaaa---atcagtagtgatgtgagataagaatgctgtgttca 408  
DB 637 acagccgatatcaaaccttattccagagatgscatttgcgaagagctgcttctcactttga 696  
OY 409 aacatgtcacttttgcgtttgctgcaagtgtgtttgtgtaaggctggaactcacc 468  
DB 697 gccctggcagatgctgtgacagatgattatgcctacgtcaataatcggaattggctg 756  
OY 469 ctgaagacactatccacgtctcctctataatgcttcagcgccagcatgttctagatg 528  
DB 757 cttaatttggagctcaacatttggccctcggttctcaccatgagggcgtggttctcagttt 816  
OY 529 gtacatctatgaccagtaaggaattgaatgctatctccaagctgggctgtgagaga 588  
DB 817 atcatcaatgacagcaaaaagtcgttgacagctacgtgacagccgctgttggcgaa 876  
OY 589 gaagctgtgcatcaatccgaacagtcagtccttaaggcccaaggagaagaactca 648  
DB 877 gagctcatgctccatcagaaatgccacagcgttggcaccacagacaaagacttggcag 936  
OY 649 aggtatacacagaatctcaagaatgcaagatgttgccataaagaagactatagctta 708  
DB 937 cagatagagctccactttagacgaagctgagaaatggggaacaaagacagattgtcag 996  
OY 709 aaagtctcttggtgctgtgtacttcttataagaagacactagactgttcttg 768  
DB 997 ggttcatatgttgccgcatgtttggccttatgtactcaactacgcttcttggcttctg 1056  
OY 769 tatggaacctctgtacttcttaatgagaacctgatatcaactcgcgtgtctgtct 828  
DB 1057 atggggttctggttccctggtatagtg-----tgacgtgatgtggtgattttcaca 1110  
OY 829 gtttcttctatgtagtaatccatagcagttatgtcatgtgagcagctccctcaactga 888  
DB 1111 gtttctatggcgcatttggatcgagatcggttcttgggaacctgtagtccaaatgtctca 1170  
OY 889 accttcgaatagccgagagcgtgctcttcatatttccagttatgataagaacc 948  
DB 1171 gcattttacaaacgctgtggccggccgacaaagatttttgaaacgattccacgtcc 1230  
OY 949 agtatagaacttctccacagctgatataaacctgaatccatagaaggaactgtgaa 1008  
DB 1231 ccatttgcatttgcatttgcgaacgaaggaagacgctgcacatttggagccacatttgag 1290  
OY 1009 tttaaaaatgttcttctcaattatccataagaacatctcataagaattctgaaagtctg 1068  
DB 1291 tttaacgaatgtcamaagcattttaaaccattcagacccgaggtcaccctcattgagatgt 1350  
OY 1069 aatctcagaattaaagctctgagagagacagtcgctgtgctgctcactaagtcagtgga 1128  
DB 1351 tctcttgcatttgccttgcgaataaacaacgctttagtgccttctgcttgcgaataa 1410  
OY 1129 agtacgtagtccagcttctgcaaggttatatgactcagatgtagtcttcatcagtgt 1188  
DB 1411 agtacggtgtgctgttggttgagcatttctacatgcctgttcgcggtatggtttgctg 1470  
OY 1189 gatgagaatgacatcagaagctttaaattgctggcatatccagaaacatatattgagtgt 1248  
DB 1471 gatggcattgacatcagaagcctcaattcctgcctgctgcacaaagatcttgggtt 1530  
OY 1249 agtcaagagcgtgttctgttcggaccacacatcagtaacaatat----- 1292  
DB 1531 agccagagagcgttcttcttttggcagcagatttataaagatttttagcagcggtctatc 1590  
OY 1293 -----caaglatgagcagatgtagtctgactgataagagatgtagagagagca 1341  
DB 1591 ggcacaaagtagaagatgaatccgagatgaagtcggggaactcattcgagaaacggcgca 1650

OY 1342 agggaaagcaaatgctgatagtattatcatgagagtttccctaataaatttaatacttgta 1401  
DB 1651 aaatggcgaatgctcatgacttatttactgcttcccttgcgaagttatagacaaatggt 1710  
OY 1402 ggggaaaaagagctcaaatgtagtggagagcagaanaacagagatgcgaattctctgtgc 1461  
DB 1711 ggccacagctggcttcttcttccatgagtgccgacaaacagcagtttcatttgcgcctg 1770  
OY 1462 ttagtgcgaaccccaagattctgattatagatgagctagctgtccctgattaga 1521  
DB 1771 gttgttagagacccaaatnctgctcctgagatgaagctacttgccttgagacaaaa 1830  
OY 1522 agcgaatcagctgttcaaatctgacgtgagaaagcgaagaagtgctgaactaacatcgtg 1581  
DB 1831 tccgaagcgtgtgattcaacgacttggagagggcagctgaaagccgaacttatttgg 1890  
OY 1582 gtaacacacagacttctactatctgcaatgacagatttgaattgacctaagaagtga 1641  
DB 1891 atgcctcatgccttcttccagatcaaaacggccacacatttgggttcgcaatggc 1950  
OY 1642 atgctgaggaagaagagacacatgctgaactaaatgycaaacagatctatatctca 1701  
DB 1951 aaatttgtagaacacaaagactcagatgaattgttgaccggcagccttattccaaa 2010  
OY 1702 ctgtgattgcaagagatttaaaaa----- 1728  
DB 2011 cttgtgagagcttcaactatcaatgaacacagaaagctgacgcttggacagcggcag 2070  
OY 1729 gctgaagacagatgtagtaatgacatatctcactgaagaagaacacactactctct 1788  
DB 2071 gctgaagatctcagatgacagatatttggcamaaatcamaactcgttcaagcagatcag 2130  
OY 1789 ctgcactctgtgaagacatcaagtcagacttcaattgacaagctgag---gaatccac 1845  
DB 2131 gatcttgacaggaaaacccacacacatttgacccgagccgacccacacagctgttccag 2190  
OY 1846 caatcfaaagagatgaagcttccctggaagctcctactatlaaaattttaaag----- 1896  
DB 2191 gcgatttctttaaagaccccccgaacacacttccgaatpactcattttagacgctgctc 2250  
OY 1897 -----ttaacaagcctgaatggccttctgtgtcttctggtgagacattgct 1941  
DB 2251 aaatttggcttcttcaaacgcccctgaatnccgacatgctcatgctgttgccttc 2210  
OY 1942 tctgtctaaatggaactgttcaatccagttatcttccatcactcttgcataaataaac 2001  
DB 2311 tcagtggttagctgtgtgctgacacacgacgaagctgcttatttcaatgacacatcag 2370  
OY 2002 a-----gttggaaataatgataaaacacatlaaagacatgagatgagaatttctc 2055  
DB 2371 acacttgcctccacgaatcacaattatgacacgttgcacatgagggatgttctgctca 2430  
OY 2056 atgatactgcatcttctgtgttcttctgtccttgcagttatctatctgcaggaattatct 2115  
DB 2431 ttatgttcttctgtgtgtgtatcatttcaagttatcagcagacacacatgctgctga 2490  
OY 2116 tacgcaagcaggggaatcttaacgatagaatlaagacacttgccttcaagcatg 2175  
DB 2491 ttggcgttatgctccagagacttattcgtgcgcgacgaagcactccttgcgagacta 2550  
OY 2176 ttatatcagatatgtgcgtgtttagtgaagaagaacagcagaagcgtctgacaaca 2235  
DB 2551 ctccggtcaagacatttcttcttgaacagagaaagaaatgacacccgctctgacactct 2610  
OY 2236 atattacatagatatagacaaatcacaagagcaagatctcagattgtcgtctta 2295  
DB 2611 ttctctgacacagacagacagacatctctcgggtttagcgttgaactgaacagatc 2670  
OY 2296 acaaaaatgcaactaacaatggaacttcaatcaatcttctatataatgagtgag 2355  
DB 2671 ttgatatacctccacgacccctgagcagcggcttatttatttgccttgcagattgggtgaa 2730  
OY 2356 atgacatctctgtatctgtatgtctgcagtaacttgcgtgtgacgaagaatgattgaac 2415

```

Db 2721 TTGGCTTAGTTTATCTCGGTGTGCGGGTTCTCTGCGCATGCGGTTTCTACCGATTC 2790
Oy 2416 gcaagaatgactgtgatttgcacaagaataagaagaacttaagcagtcygaagata 2475
Db 2791 TATATGCTAGGCCATCTTCATCAACGCTCCACGTTGCTTATGAGGATGCTCAACTTT 2850
Oy 2476 gcaactgagcttggagaataatactactagtgctactaagaaggaaagaccttc 2535
Db 2851 GCTTCCGAGGCTACATCGTCTATCCGCACAGTTGGCTCTTTAACCCGGGAAAGGATGTC 2910
Oy 2536 gagcaaatglatgaagagatgcttcaagctcaacacagaataactcgaagaagacag 2595
Db 2911 TGGGAGATTTCACATGCCAGCTTGACGACAAAGCAGAGACCAAGCTATCTCTGCTTG 2970
Oy 2596 attattgaaagctgttataagcatcagcagccttataatttgcctatgcaagagg 2655
Db 2971 AGGTCACTCCCTGTATATCGCTGTCGCAAGGACACTTGTTCCTGCGGTGCGCTCGG 3030
Oy 2656 ttctgattgagcctatttaataagctggaagcgaatgagcccgagaggcagcttcata 2715
Db 3031 TTTTGTGACGAGGAGACACTTCTTGTCACACAGATAGACATTTTCGCTTCTTTGTT 3090
Oy 2716 gttttactgcaattgcatatgagctatgcccacggaagaaacgctglttggcct 2775
Db 3091 TGTTCTCGAGATTCTCTTTGCTGCTCAATCCGCGGACCGCTTTCTTGCACACA 3150
Oy 2776 gaatttccaagccaataatcgaggctgagcctgctgttgccttcttggaaagaacca 2835
Db 3151 GACATGGGCGAGAGAGATCGCGCCGCGCAATTCGACGACTGTTCACCGAAAGCCA 3210
Oy 2836 aatatagacagccgagctcaagaagaagaaagcagacacatgtaaggaattagag 2895
Db 3211 CAAATTTGATTAATCGTCTCAAGAGCCGAGAAAGCTCGAAACGCTGGAAGTAAATCGAA 3270
Oy 2896 ttctgagaagctcttcttctcaatgctgcagagatglttcaatcccttcgtygcta 2955
Db 3271 TTTAGGAAGCTGACTTTCAGATACCGGACCGCCGCAAGACAGCTGTCTGCGGCGCTG 3330
Oy 2956 tccctcaagtattgagcgaaggaagaagcagtagcattgtgaggagcagcgctgtggaaa 3015
Db 3331 GACCTGACCGGTAAAGCTGAGCAATATGTGGCTTGTGGACCAAGCGCTGTGGCAAG 3390
Oy 3016 agcacttctgttcaacttctgcaagagactttagaccccgtagcaaggaagtgctggt 3075
Db 3391 AGTACACCATTTGCATTTGCTTGAAGCCTTTTACGATGCGATGCGGCTTCTTGT 3450
Oy 3076 gatgtgtgagatgcaaaagaattgcatagcagtgctccgttcccaaatagcaatcgt 3135
Db 3451 GATGGGAAGGACATAGTAACTAATATCAACTCTACCGCAGCTTCTGTCACATGCTC 3510
Oy 3136 cctcaagagcctgtgcttctcaacgtagcagcattgctgagaacaatgcccatagtgaac 3195
Db 3511 AGCCAGGAGCGCACACTGTACCAAGGCGCACATCAAGAAACATCTTACTTGTATTTGTC 3570
Oy 3196 agccgtgtgtgcccattatagatgaagaagccgcaatgcaagcaatltccattct 3255
Db 3571 GAAGATGAGCTACCGGAAGAAATTTTGATTAAAGCTTGCAAGAGCCTAATATCTACGAC 3630
Oy 3256 ttattgaaagctctcccttgaataatacacacacaagtlgagcaggaagagacagctt 3315
Db 3631 TTCAATCATCTGCTCCCGGAGGCTTTTATACAGTTGTGGCAGCAAGGAGGACATGTTG 3690
Oy 3316 tctgagcgcagcaaaagaactagctatgcaagggcctcttccaaaacccaat 3375
Db 3691 TCTGGGGGCAAAAGCAAGCTGTGGCAATTTGCCGAGCCCTTCTTGGGATCCCAAAATC 3750
Oy 3376 ttattgtgagagagccctcagccctcgaataagagagagagagtggtttagcat 3435
Db 3751 CTTTCTTCTGATGAAAGCAGCTGACCTCGACTCCGAACTCAAGAAAGGTCCTCCAGGCG 3810
Oy 3436 gcccttgaataagcagagcaggaagagatgcctagtgtagtcaactcagagctctctga 3495

```

```

Db 3811 GCTTGTAGTCCGCTGCCCGGAGCGCAACCAATCGCCTTGGACACCGACTGACGACG 3870
Oy 3496 attcaagcagagattgatagtgcttgcacaatggaaagataaagaaaggaact 3555
Db 3871 ATTCAAAAGCGGAGGATTAATGTTTTCGACCAAGCAAGATCTGTGAAGGGAACG 3930
Oy 3556 catcaagctcctcgtgagaatcagacatatatttaagtagtgaatgacag 3609
Db 3931 CACACGAACTGCTCCAGAAAAGGCGGCTACATCAAGCTGTGTCACACTTCAG 3984

RESULT 13
US-08-996-545-3
: Sequence 3, Application US/08996545
: Patent No. 5928898
:
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Waard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of
: TITLE OF INVENTION: Aspergillus nidulans
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/996,545
: FILING DATE:
: CLASSIFICATION: 435
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11766
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
:
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4002 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLBCULE TYPE: mRNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: US-08-996-545-3

Query Match 13.7%; Score 496.8; DB 2; Length 4002;
Best Local Similarity 34.7%; Pred. No. 9.7e-14;
Matches 1226; Conservative 479; Mismatches 1727; Indels 102; Gaps 7;
Oy 172 cgtatattagttggaatagtgcttgcctgctgatttggttgcatacaagattccttg 231
Db 457 CUGUACUUCGUDUACUCCUGUACGCGGAGUUGUCACUGUCUAGUAGUACUGUUGGC 516
Oy 232 tggatataactgcagcagacagacaaagagatlcgaaacaglttttcaatcagtt 291
Db 517 UUCACUUAACCGGAGAAACGCAAGCAGAGAGUCCGCGAGUUAUUDACUUAAGUCUAVC 576
Oy 292 ttggaacagagacatggctgtttagtagctgtgacatcgtgtgaacttaacactgcag 351
Db 577 CUGGCCAGAAAUUGGCUAUAUUGAUAUAUCUGUGCGGCGGAGAGUGACCCGUAUA 636

```





Db 2791 UAUAGCAGCCAGAUUCAAUACAGCUCCAAAGCUUGCUUUGAGGAGAUUCCAGAAACUUU 2850  
Oy 2476 gcaacttgaaagcttggagaataatactactagtctattacaagaagaaagccttc 2535  
Db 2851 GCUUGCGGCGACAUUCCUUAUCCGACAGUUGCGCUUAUACCCGGGAGGAGUUGC 2910  
Oy 2536 gagcaaatgatatgaagatgcttcagactcaacagaataactcgagaagaacacag 2595  
Db 2911 UGGGAGAUUUUOCCAGCCAGCUUUAAGCAGAGCAGACAGCUUAUUCUUCUUG 2970  
Oy 2596 attatggaagctgttatgcatctcagcatgacctatataatttgcctatgcagcag 2655  
Db 2971 AGGUDAUCCCGUUAUAGCGGCGGACAGGACAUUUGUUGCGUUGCGCUGCGG 3030  
Oy 2656 ttctgattgagcctatttaattcaactgtagcgaatgaccccgaaagcgatgtcata 2715  
Db 3031 UUUUGGUAAGGAGACAUUUCUUGGUCACAGAGUAUAGAUUUUCCGCUUCUUGU 3090  
Oy 2716 gtttactgcaattgcatatgagctatgagcctcgaaaaagcgtctgttgcctcct 2775  
Db 3091 UGUUUCUCCGAGAUUUCUUGUGUCUCAUCCGGGCGACCGUUCUUCUUGCAGCA 3150  
Oy 2776 gaataatccaaagcgaatcgagggtcgagcatctgttgcctgttggaaaaagaca 2835  
Db 3151 GACAUGGCGCAAGGAGAAUAGCGGCGCCGCAUUCGACGACUUGUCCAGCAAGGCA 3210  
Oy 2836 aatatagcagcgcgagcagcaagaagaagaaagcagacacatgtgaaggaatttag 2895  
Db 3211 CAUAUAGUAUACUGGUCUAGAGAGGCGAGAGGCGAAACGUGGUAAGUUAUACGAA 3270  
Oy 2896 ttctgagaagctctctctctatcatcgtcgccagatgtttcattcctcgttgccta 2955  
Db 3271 UUUAGGAACGUGACAUUCCAGAACCCGACCGCCAGAACGCUUGUCUGCGCGCU 3330  
Oy 2956 tccctcagatattgagcaggaagaacagtagcatttggggagcagcgtgtgtgaaa 3015  
Db 3331 GACCGGACCGUAGACCGGAGCAUAUUGGCGUUGCGGACCCGCGUUGUGGCAAG 3390  
Oy 3016 agcaacttcttcaactctctgagagacttataagcccgtagcagaagaagtgtgtt 3075  
Db 3391 AGUACACCAUUGCAUUGGUCUAGCGCUUUAUAGAGUUGCGGCGCCUCCAUUCUUGU 3450  
Oy 3076 gatgtgtgtgctgcaaaagaaatgatactagtgctcgtctcccaataagcaatcgt 3135  
Db 3451 GAUGGGAAGGACAUUAUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3510  
Oy 3136 cctcaagagcctgtgtctctcaactgagcagcatctgtagaacaatcgctatgtgaca 3195  
Db 3511 AGCCAGGAGCGACACUUGUACAGGCGACCAUAGGAAACAUUUAUUGUUAUUGU 3570  
Oy 3196 agccgtgtgtgtgcaattgatagataagaagcgaagcgaatgacgaataatcattct 3255  
Db 3571 GAAAGUAGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3630  
Oy 3256 ttatitgaaggtctctcctgagaataatacaacacagaatltgactgaagaagcagact 3315  
Db 3631 UUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3690  
Oy 3316 tctgtgagcagaaagaaagactagatttgaagagctctctccaaaaacccaanaat 3375  
Db 3691 UCUGCGGCGCAAAACAGUUGGCGCAUUGCGGCGCCUUCUUGGAGUUCGCAAAUUC 3750  
Oy 3376 ttatgtgtgtgtgagcacttcaagccttgataatgacaggaagaaggtgttagcat 3435  
Db 3751 CUUUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3810  
Oy 3436 gcccttgataaagcaggaagcaggaagcagcatcctagtgtcactcagaagcgtcctgca 3495  
Db 3811 GCUUUGAUAUUGCGGCGGAGCGGAGCAUUAUUAUUAUUAUUAUUAUUAUUAU 3870  
Oy 3496 attcagaacgagattgtatgtgttctgcaaatgaaagataagaagaagaact 3555

Db 3871 AUUCAAAGGCGAGCGUUAUUAUUGUUGGACCAAGGCAAGUCCGUGAAAGCGGAACG 3930  
Oy 3556 catcaagagctccttgagaatcgagacatatatttaagttagtgatgcacag 3609  
Db 3931 CACAGCGAUCUGGUCAGAAAAAGGCGCGUAUUAUUAUUAUUAUUAUUAUUA 3984

RESULT 14  
US-09-328-320-1  
Sequence 1, Application US/09328320  
Patent No. 6228615  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Waard, Maarten A.  
APPLICANT: Peery, Robert B.  
APPLICANT: Andrade, Alan C.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atpD of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/328,320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/996,545  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872  
REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..4002  
US-09-328-320-1

Query Match 13.7%; Score 496.8; DB 4; Length 4002;  
Best Local Similarity 48.2%; Pred. No. 9,7e-141;  
Matches 1705; Conservative 0; Mismatches 1727; Indels 102; Gaps 7;

Oy 172 ctgtatattgttgtaaatgagtggtgtgtccttgatttttgattacatacaattccttg 231  
Db 457 CTGTACTCTGATATACCTGATCGGATCGGCGATTGTGCACTGTATGATGATGATG 516  
Oy 222 tggatataactgacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 291  
Db 517 TTCACTTATACCGGAGAAACAGCCACGACGAGAAAGATCCGAGTATTACTTGAGTCTATC 576  
Oy 292 ttgacacagacatcgctgtgttgatagctgtgacatcgttgaaactaactcgcacatg 351

```
Db 577 CTGGCCGAACATTTGGCTATTTTGATTAACCTCGTGCCGGGGAAGTACACACCCGATATA 636
Qy 352 acagacattgacaaa---atcaatgtagtcatctgagataagaattgcctgcttcaaa 408
Db 637 ACACCCGATACAAACCTTATCCAGGATGGCATTTTCGGAGAGGTGCGCTCATCTTGGACT 696
Qy 409 aacatgctacttttctgattgagctgagctgagttggttggtaaggagctgaaactcacc 468
Db 697 GCCCTGGGACATTCGTACAGCATTCATTTAGCCCTACGTCGAATATGTAAGTTGGCT 756
Qy 469 ctaatgactatccacagctccctctataatgcttcaagcgacagctgcttcaagatg 528
Db 757 CTATTTTGGACACTCAACATTTGTGGCCCTCGTCTCAACCATAGGGCGGTGTTCTCAGTTT 816
Qy 529 gtcaatcctatgaccagtaagaaatlaagtgctatctatccaaagctgaggctgagcaaa 588
Db 817 ATCATCAAGTACACGAAAAAGTCCTTGACAGCTACGTCAGCGGCGGCACTGTTGCGGAA 876
Qy 589 gaatctctgcatcaatccgacagtaatagcctttagggcccgagagaaagaacttcaa 648
Db 877 GAGGCTACAGCTCATCAGTAAGTAATGCCACAGCGTTTGGACCCAAAGACACTTGGCAAG 936
Qy 649 aagtaacacagaaatcacaagatgcaaaaggaatttggcatataaaaggaactatagctca 708
Db 937 CAGTATGAGGTCCACTTAGACGAAAGCTGAGAAATGGGAGACAAAGAACCATTTGTCTATG 996
Qy 709 aaagtgtcctctggtgctgtagtacttctatgaatgaaacctatgagcttgccttttgg 768
Db 997 GGTTCATGATTTGGCGGCGCTTTGGCCCTTATGTACTGCAATCAGGCTGTTGGCTTTCGG 1056
Qy 769 tatgaacctccttgatcttaatgaaagaaacctgatatataccatcgagacgttcttgct 828
Db 1057 ATGGGTTCTCGTTTCTGCTGGTAGATG6-----TGCAGTGATGTTGGATATTCACAA 1110
Qy 829 gtttctcttagtgaatccatcagacagttatgcatltgagcagcagctccctcaactttaa 888
Db 1111 GTTCTCATGGCCATCTTATGATGGATGCTTTCCTTTGGGGAAGCTTAGTCCAAATCTCA 1170
Qy 889 acccttcgaatagcccgagagagctgcttcatatatttccagatgatgataagaacc 948
Db 1171 GCATTTTCAAAACGCTGTGGCGGCGGCAAAAGATATTTGGAACGATGATGCGCAGTCC 1230
Qy 949 agtatagataacttttccacagctgagataaaacctgaatccatagaagaagacgttggaa 1008
Db 1231 CCATTTAGATTCATATTCGAACGAAGGGAAGACGCTCGACACTTTTGAAGGCGCACATTGAG 1290
Qy 1009 ttaaaaaatgttcttcaatataatcacaagacacatcatcaagaattctgaaagctg 1068
Db 1291 TTACGCAATGCAAGCATATTTTACCATCTAGACCCGAGGTCCACCGTATGAGAGATGTT 1350
Qy 1069 aactcagaatlaagctcagagagacagtcgcttgctgagttctcaatgagcagttggagaa 1128
Db 1351 TCTCTGTCAATGAGCCCGCTGGAAGAAAAACACGCTTAGTCGCGCCCTCTGGCTCTGAAAA 1410
Qy 1129 agtcagagagtcceagctcttcagagagttatagatccgagatgagcttatacagtg 1188
Db 1411 AGTAGGGGTGTCGGCTGTGGTGAAGGATTTTACATGCTGTTCGGGTACGGTTTTGTG 1470
Qy 1189 gatagaatgacatcaagctttaaagtcgagcatctagacacatattgagatgagtg 1248
Db 1471 GATGGCCATGACATCAAGAGACTCAATCTCCGTGGCTTCCGCCAACATATCTTTGGTT 1530
Qy 1249 agtcaagagcctgttctgctgaggaacacatcagtaacaatat----- 1292
Db 1531 AGCCAGAGGCTGTCTTTTGGCACGACGATTTATGAATAATATGACGACGGTCTCATC 1590
Qy 1293 -----caagtatgagcagagatgagtgtgactgataagagatgagagagca 1341
Db 1591 GGCACAAAGTACGAGAAATGATCCAGAGATGAAGTCCGGGAATCATGACGAACCGGCA 1650
Qy 1342 agggaagaatgcatgattatcatcagtgagttcctaataattatcatctgta 1401
Db 1651 AAATGGCAATGCTCATGACTTTATTTACTGCTTGCCCTGAAGGTTATGAGCAATGTT 1710
Qy 1402 ggggaaaaaggaagctcaaatgagtggagggagaaacagagatcgcgaattgcttgctc 1461
Db 1711 GGGGAGCGTGCTTCTCTCTTCAAGTGGGCCAAGAAACGCGCATTTGCAATGCGCCGTGCC 1770
Qy 1462 ttaattcgaaccccaagattctgattttagatgagagtagctgcttgccttgatcagaa 1521
Db 1771 GTTGTATTGACCCCAAAATCTTGCTGCTGATGAGTACTTGGCTTGACACAAAA 1830
Qy 1522 agcaatgacgtgttcaagctgcacactggaagaagcgaaagtgctgactacaatcgtg 1581
Db 1831 TCCGAAGCGGTGTTCAAGCAGCTTTGGAGAGGCGAGCTGAAGCGCGGAATCATATTG 1890
Qy 1582 gtagcaacggaacttctactatctcgaatgagtgcaagattgattgacccaaagatgta 1641
Db 1891 ATCGCTCATGCGCTTTCCACCATCAAAACGGCGCAACATATGAGTGTCTGTCAATGGC 1950
Qy 1642 atgctgagggagaaagagacatgctgactaatggaacaaacgagctatattatca 1701
Db 1951 AAATTTGCTGAACAAAGAACTCAGCATGAATTTGTTGACCGCGAGCGGCTTATCGCAA 2010
Qy 1702 ctgtgatlgtcacaggaalatataaaaa----- 1728
Db 2011 CTGTGAGAGGCTCAACGTATCATGATGACAGAAAGAAAGCTGACGCTTGAGAGCGCGAC 2070
Qy 1729 gctgataacagatgagtgatcaatgacatatctactgaaagaagaaaccaactcacttct 1788
Db 2071 GCTGAGATCTCCACCAAAAGCAGATATTTGCCAAATAAACAACCTGCTCAAGCGCATATCC 2130
Qy 1789 ctgacactgtaagaagacatcaagctcagactatctgacaagagctgag---gaatcacac 1845
Db 2131 GATCTCGACGGAAACCCACAAACCATTTGACCGCACGGGCGCACACCAAGTCTTTCACAC 2190
Qy 1846 caatcacaagataagcttctcctgaagctctcctatataaaatlttaag----- 1896
Db 2191 GCGATTTCTTTAAAGACCCCGCAAAACACTCGGAATACTATTTGACGCGTGCCTC 2250
Qy 1897 -----ttaacaagcctgaatgagccttctgtgcttctgagcttgcagatgagct 1941
Db 2251 AAATTTGTTGCTTCTCTTACACCGGCTCGAATAATCCGTACATGCTCATGCTTGTCTTC 2310
Qy 1942 tctgttctaagtgaagacgtgtcactcagatatttccatcatcttgcgaataatlaaac 2001
Db 2311 TCAGTGTAGCTGTGGTGCGCAACCCACGCAAGAGATGCTATATGCTTAAGCCATCAGC 2370
Qy 2002 a-----tgttggaataatgataaaacacatlaaagcatgataagaaattatctc 2055
Db 2371 ACATCTCGGCTCCGAAATACAAATATAGCAAGCTTTCGACATGATGCGGATTTTGATCA 2430
Qy 2056 atgatatcgtcaatttgggtgttatttgccttgcagttatlttcaagagatlaatt 2115
Db 2431 TTGATGTTCTTCTGTTGTTGATCATTTACAGCTTACACGCACTCAACCAATGCTCTGCA 2490
Qy 2116 taagcgagaagggaaattttaaagatgataaagaacattgagccttcaaaagcaat 2175
Db 2491 TTTGCGGTATGCTCCGAGAGACTTATTTGCTCGCGGAGAGACATGCTTTTGCGACGATA 2550
Qy 2176 ttatatacaggaatattgcttggatttgaatgaagaagaaacagacagagagcttgaaca 2235
Db 2551 CTCGCTCAAGACATTTGTTTGTGACAAAGAAAGAAATAGACACGGGCGCTTGACCTCT 2610
Qy 2236 atattaacatagatatagcaacaaltcaagagagcaacaggttccaagattggcgctcta 2295
Db 2611 TTCTGCTCACCGAGACCAAGCATCTTCCGCTGTAGCGGAGTGACATCTAGACGACGATC 2670
Qy 2296 acaaaaaatgcaactcaaatgaggaacttcagttatcatcttccattatataatgagtgag 2355
Db 2671 TTGATGACCTCCACGACCTTAGAGCGGCTTATCATTTATTTGCCCTTGCGATTTGGTGAAA 2730
Qy 2356 atgacatcctgattctgattatgctcctcagttacttgcgtgagacagaaatgaltgaacc 2415
Db 2731 TTTGGCTTAGTTTGTATCTCGTTGTGCGGTTTCCCGGATGCGGATGCGGTTTACCGATTC 2790
```

OY	2416	gcagcaatgactggaatttgcacaaagaataagcaagaactcaagcatgcttgyaagata	2475
Db	2791	TATATGCTAGCCCAATCTTCATACAGCCTTCAAGCTTGTATGAGGATCTGCAAACTTT	2850
OY	2476	gaaactggaagctttgagagataaagtaactatagtgtcaataagaaggaaagaaagcttc	2535
Db	2851	GCTTGGAGAGCTACATCGTCTATCCGCACAGTTGGCTATTAACCCGGGAAAGGATGTC	2910
OY	2536	gaacaaatglatgagagatgtcttcaagactcaacacagaatacctcgaaagaagacag	2595
Db	2911	TGGGAGATTATTACCATCCAGCTTGACGCAAGGACGAGCACAGTAACTCTGTCTTG	2970
OY	2596	atatttgaagactgttatcatctcaagccatgtcccttatatatattttgctatbgaagag	2655
Db	2971	AGGTCAATCCCTTTATATATGCTGTGTCGAGGCACTTTGCTTCTTGCGCTTGCGTCGGG	3030
OY	2656	tttcgaatttggagaccatttaattcaagcttgcagatgacccacagagggacglttcata	2715
Db	3031	TTTTGGTACGGAGGACACTTCTTGGTCAACACGAGTATGACATTTTTCGCTCTTTGTT	3090
OY	2716	gtttttactgtcaatttgcataatggagctatgagccatcgcagaanaaagcgttcglttgcctc	2775
Db	3091	TGTTTCTCCGAGATTTCTTTTGGTGTCTAATCCGGCGGACCGCTCTTTTCTTGACACA	3150
OY	2776	gaatatcccaagcaaatcgcggggtcgcgcatctgttttgccttgttggaaagaaccca	2835
Db	3151	GACATGGGCAAGCGGAAAGATATGGCGCGCGGAATTCGACGACTGTTCGACCGCAAAACCA	3210
OY	2836	aatatagacagccgcagctlaaagaaggaaanaagcacagacatctgttaagaagaattagag	2895
Db	3211	CAATTTGATTAATCTGGTCTGAAAGGGCGAGAACCTCGAANCGTGGAGGTGAATCGAA	3270
OY	2896	tttcggaagctctcttcttctctatccatgtcgcagaatgttttcaactcctcgttggcta	2955
Db	3271	TTTATGGAAGGTGACTTTCATGATACCGGACCGCGCCAGAACAGCCTTGCTGGCGGGCTTG	3330
OY	2956	tcacctagatttgaacggaagaaagaacagtcagatcttgttgggaagcaagcgttgggaaa	3015
Db	3331	GACCTTACCGGTGAAGGCTGTGCACATATGTTGCGCTTGTGCGACCCACACGGTGTGGGACG	3390
OY	3016	agacctcttgctaacctcttcgcgagaacttataagaccgctgcaagagacaagtgcgttt	3075
Db	3391	AGTACCAACCTTCATTTGCTTGAGCGCTTTTACGATCGATTTGCCGGGTCCATCCTGTT	3450
OY	3076	gatgtgttgaatgcaaaagaatttgaatgtaacgttgcctccgttcccaaatagaatcgtt	3135
Db	3451	GATGGGAAGSACATTAAGTAACTAATATATCACTACCGCAGCTTCTGTGTACTGTGTC	3510
OY	3136	cctcaagaagcctgtgtccttccaactcagcatctgctbgaagaacatcgctatggtgacaac	3195
Db	3511	AGCCAGAGAGCGACACTGTGTCCAGGGCCACCATCAAGGAAAAACATCTTACTGTGATTGTC	3570
OY	3196	agccgtgtgtgtgcattagatgataaagaagccgcaaatgtagaataatcatcttc	3255
Db	3571	GAAGATGACTTACCGGAGAATTTCTTGATTAAAGCCTTGCAAGGACGCTTAATATTTAGAC	3630
OY	3256	tttattgaaagcttcccttggaaataacaacacaagatttgaactgaaagagacagctt	3315
Db	3631	TTTATCATGTCCGCTCCCGGAGGCGCTTTAATATACAGATTGTTGGCACACAAAGGAGCATGTG	3690
OY	3316	tctgtgcggccaagaacaaagactagctatgtgcaaggctccttctccaaaaaaccanaatt	3375
Db	3691	TTGTGGGGGCCAAAAGCAAGCTGTGTGGCATTTGCCGAGCCCTTTTTCGGGATCCCAAAATC	3750
OY	3376	ttatgttgaatgaagccaacttaagccctcagataatgacatgtgaagatgaaaggtgttgcagcat	3435
Db	3751	CTTCTTCTTCATGATGAGGAGCTGACGCCCTCGACATCCAGTCTAGAAAAGGTGTCAGGCG	3810
OY	3436	gaccttataaagccaagcaagcgggaagacatgacctagtgtgcatacaagcctctcttgcga	3495
Db	3811	GCCTTGGATGCGCTGTGCCAGGCGCGGAACCAATTCGCCGTTTGACACACGACTACGACGC	3870

```

Oy 3496 atcagaacgcgaattgatgtgtgttcgtacacatggaaagataaagaaacgaagact 3555
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3871 ATTCAAGGCGGAGCTATCTATCTTTTCAGACCAAGGCAAGATCTTGAAAGCGGAGC 3930
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 3556 catcaagccctcagaacatcgcagacatatatttaagttagtcagacag 3609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3931 CACAGCGAAGTGTCTCAGAAAAAGGCGCGTACTACGAGCTGTCTCACTTGACG 3984

RESULT 15
US-09-328-320-3
: Sequence 3, Application US/09328320
: Patent No. 6228615
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Maard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of
: TITLE OF INVENTION: Aspergillus Nidulans
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/328,320
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/996,545
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11766
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4002 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: mRNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
US-09-328-320-3

Query Match 13.7%; Score 496.8; DB 4; Length 4002;
Best Local Similarity 34.7%; Pred. No. 9,7e-141;
Matches 1226; Conservative 479; Mismatches 1727; Indels 102; Gaps 7

Oy 292 ttggcagagacatcgctggttttgatagctgtgacatcgggtgaacttaacactgcacg 351
      :||| ||||| ||||| :||| ||||| ||||| ||||| ||||| ||||| |||||

```

D	b	577	CUGGCGCCAGAACAUUGGCUUUUUUGAUAAACUUGGCCGGGGGAAGUGACACACCCGUAUA	636
O	y	352	acagacacttgacaa---atcag tga tgg lat tga gataa gata gct tgc tgc ttt caa	408
D	b	637	ACAGCCGGAUACAAACCUUUAUCCAGANUGGCAUUNUGGAAAGAGUGGUCUACUUAUCU	696
O	y	409	aacatgtctacttcttgcattgcctgtgcagttggtttgtgtgaagggtcgtgaactcacc	468
D	b	697	GCCUGGCGACAUUCCGUGAGCAGCAUUDCAUUAUUGCCUACGUCUAUUAUUCUGGAGAGUGGCU	756
O	y	469	ctagtgaactatcatcagctctcctcttataagctcgaagcgcaagatgltctagatg	528
D	b	757	CUAAUUUUGCAACUUCAAAUUUUGUGGCCUUCGUUCCACCAUUGGGCGGUGUGUCUUCAGUU	816
O	y	529	gtcatctcatatgacacag taa gaa gataa g tgc ctat tcca aag cgt ggg gct tgc tga a	588
D	b	817	AUCAAUCAAUUAACGCAAAAAGUGGCUUUGAACGUAUAGUGGCAAGGGGCGACUGUUGCGAA	876
O	y	589	gaagltctgtcatcatalccgaacagcatcatagcctttaggtggccagagaaagaaactcaa	648
D	b	877	GAGGUCACUACGUCUCCACUACAAAUUGCCACAGGUGUUGGCAACCAAGAACGUUGCAAG	936
O	y	649	aggtatcacacagaaatcccaagaagttgcaagaagtttggcatabaaagaactatagcttca	708
D	b	937	CAGUAUAGAGUCCAUUAGCGAAGGUGUGAUAUUGGGGAACAAACACAGAUUGUCUAUG	996
O	y	709	aagtgctctctgtgtgtgtgtactcttataagaagaaactatgagactgttcctttgtg	768
D	b	997	GGUUCCAGAUUUGGCGCCAGUUUUUGGCCUUAUGUACUCCAGACUACGUGUCUGGUCUGG	1056
O	y	769	tatgaaacctctgtatcttcttaatggagaagctgtgatalaccatcgtggagtgttcgt	828
D	b	1057	AUGGUGUUCUGUUUCUGUAGAUUG----UGCAGUGCAUUGUGGUGAUUUUCUCA	1110
O	y	829	gtttctcttagtgaatccatcctagcagttatgtcatgtgagcagcagtcctcacttgaa	888
D	b	1111	GUUUCUACUUGGCCAUCUUGAUCGGAUUGUUCUUCUUGGGAGACGUUAUUCAAAUUUCUCA	1170
O	y	889	accttgcgaatagcccgaggagctgtccttcatattcttcaggttatgataagaagacc	948
D	b	1171	GCANUUAACAAACGUGUGGCGCGCGCCGCAAGAUUUUUUGAACAAGUACGUAUUGCCAGUCC	1230
O	y	949	agtatagataacttcttccacagcttggatataaaacctgaaatccatagaaagaaacttgaa	1008
D	b	1231	CCAUUAGAUCCAUUUUUGAAGCAAGAAAGGAAGACGUCUACCAUUUUUUGAGGCGCACAUUUGAG	1290
O	y	1009	tttataaatgttcttcttccatattcatcatcaagaagccatcatcaagatcttgaaggtctg	1068
D	b	1291	UUACGCAAUUGUCAAGCAUUAUUUCCCAUUCUAGACCGGAGUACCGUUAUUGAGAUUU	1350
O	y	1069	aatcccaagaattgaatctgtgagagacagtcgcctgtgtcgtcctaagtgagcgttggag	1128
D	b	1351	UUCUUGUCAAUUGCCCGGUGGAAAAACACCGCUUUUAUGUGCGCCUUCUGGCUUGGAAAA	1410
O	y	1129	agtaaggtagtccagctctcgcagaggtatataatgacgtgaaatgagttcatatgctg	1188
D	b	1411	AGUACGGUGGUGCGCUUUGGUGAGCGAUUUCUAUUGCCUUGCGGUAUGGCUUUUGUG	1470
O	y	1189	gatgtagaatgacatcagacgtttaaatgtgcgcatlatacgagaccatattggaagtgct	1248
D	b	1471	GAUUGCCCAUUGCAUCAAAGGACUUAUUCUUGGUGGCUUGGCCAACAGAUUUUUUGGCU	1530
O	y	1249	agtaagaagcctgttgtgtgtcggagaccacatcagtaaaatat-----	1292
D	b	1531	AGCCAGAGCCUUGUUUUUGGCGACGAGAUUUUAACAAAUUUAUAGCACGCGUCCUAUC	1590
O	y	1293	-----caagtatggacagatgatatgtactgtacgtgaagagatgagagagacaa	1341
D	b	1591	GGCAACAAGUACGAGAUUAGAUUCCGAGAUUAAGUUGCCGGGAACUUCAGAACACCGCGCA	1650
O	y	1342	agggaaagcaaatcgttatgtatttcatatcagtagagtttccataaatttaatacatgtgta	1401
D	b	1651	AAAUUGCGAUUGUCUUAUGCAUUUAUUCUGCCUUGCGGAGGAGUUUAUGAGCAACAUUU	1710

QY	1402	gqggaaaaagagctcaatgagtgtgaaggycagaacacagagatgcgaattgctctgtgc	1461
Db	1711	ggcgagccggccuuuuccuuuacagggggccaaacacacggauuacauuucccguc	1770
QY	1462	ttagttgaaaccccaagatttcgattttagagtgagctgcagctgcgccttgattcga	1521
Db	1771	guguuuagucacccaaaauucucucucugauaagcuaucuuccgcccugacacaaa	1830
QY	1522	agcaatgcagctgttcctcaagctgcagctggaagggagcaaaagttcgaactcaatg	1581
Db	1831	uoccgaaagccgguguucaacgacgucuuagacggcgacgaaagcccaacuaucuu	1890
QY	1582	gtagagccgaacttcttactatttcgaagtgagcagatttgattgtgacctaaagta	1641
Db	1891	auccgucuuucccuccuuccacgaauaagcgccacacauuugucuuuucgcauu	1950
QY	1642	atgctgcggagaaaggacacatgctgaaactaigtgcacaaacgagttacatattca	1701
Db	1951	aaaauuucguaaacaaaggacucacgaauuuguuagaccggagcgccuuauucgaa	2010
QY	1702	cttgtatgtcaacaggaatataaaaa	1728
Db	2011	cuuaggagcucacacauuacauaagaaacagaaagacugacccuugagcgccgac	2070
QY	1729	gctgatacagatgagtgagtgcaatgacatctctacgaaagaagacaactcctgc	1788
Db	2071	gcuaagagauccacgaauucgaaauuuccaaaaucaaacuuccgucacagccauuc	2130
QY	1789	ctgcactctgtgaaagagcatcaagctcagacttcattgacaagagctgag--gaatca	1845
Db	2131	gaucucgacggaaaaacccacacacauuugacggcagcgccacccacaaucuuuuc	2190
QY	1846	caacttaagagataagttcttcctggaagttcttcttaataatttaaag-----	1896
Db	2191	ggcauucuuuuaaaagaccccccgaaacacauccgaaauuacuaauuagacgucuc	2250
QY	1897	-----ttaacaagccgaaatggccttctgtgtctcgtgagacatggct	1941
Db	2251	aaauuuuguuucuuuacacccgcuuagaaucgguacugucuuugucuuuc	2310
QY	1942	tctgttctaataatggaactgtgtcatccagtaatttccatcactcttgcacaaataa	2001
Db	2311	ucaguguaucgugugggccacacccacgaacgucuaauuucuaaacccuuaagc	2370
QY	2002	a-----tgttggaataatgataaaccacataaagcatgagtgcgaatatctcc	2055
Db	2371	acacucucgucuccacgaauacaaauuagacacucucgaacuaugucgaauuucg	2430
QY	2056	atgatatcgtcaatttgggtgttatttgccttgcagttatttcaatgcaggattatt	2115
Db	2431	uugauugucucguguguuuaduaucagauuaduaacagacuaacaaauugucua	2490
QY	2116	taaggcagaagagggaattttagatgagttaagagacttgagccttcaagcagt	2175
Db	2491	uuuugccgaugucccgagacacuaauuugcggcgccgagacacucuccuuucgac	2550
QY	2176	ttatacagagatattgcctggtttgtatgataaagaaacagacacaggaagcttga	2235
Db	2551	cucgcucaagacauuugcuuuuugacaagaaagaaugacacggcgccucucgac	2610
QY	2236	atattagcaatagatatagcacaatttcaagagcaacaggttccagaattgagctta	2295
Db	2611	uuccgucacacgagacgaacacuaucuuucgguguuagcgugucuaucuaagc	2670
QY	2296	acacaaatgacataacatgagagcttgcattacatttccattatagatgagag	2355
Db	2671	uuagaaacucacagaccuuaagaaagcgcucauauuuuucccugccauuugugua	2730
QY	2356	atgacatctcgtattctgagtattgtctcagtaactgtccgtgacaggaattgaaac	2415
Db	2731	uuagccuaguuuuuauucguguuuuccgucuuucgacaaugcguguuuucacgauc	2790

```
OY 2416 gcaacatgactgattgtccacaagaataagcaagaacthaagcatgctggaagaata 2475
Db 2791 UUAUUGGUCAGCCCAAGCUCUACAGCUCUACAGCUCUUAUAGAGGAGUACCAAAUUCU 2850
OY 2476 gcaactgagcttctggaataatactagtactagtctattcaacagggaagaacgtctc 2535
Db 2851 GCUUCCGGGACACAUCCUUAUCCGACAGUUGGUCUAUUAACCGGGAAGGAGUUGC 2910
OY 2536 gagcaaatgatatgaagagatgctctcagactcaacacagaataactctggaagaacag 2595
Db 2911 UGGGAGAUUUAUACAGCCAGCUCUACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2970
OY 2596 attatggaagcgtgtaatgcatctcagcagtcgcttataataatttgcctatgcaagag 2655
Db 2971 AGGUCAUCCUUGUUAUUGCGUGCGAGGACAUUUGUUCUUGCGUUGCGUGCGG 3030
OY 2656 ttctgattgagcctatttaattcaactgagacgaatgaccccaagagcagtctcata 2715
Db 3031 UUUUGGUAAGGAGGACAUUUCUUGGUCACGACAGUAUAGACAUUUCGCUUCUUGU 3090
OY 2716 gtttactgcaattgcatatggaagctatgagcatcggaagaagcgtcgtttgctcct 2775
Db 3091 UGUUUUCCGAGAUUUUUCUUGGUCUACAUCCGGGACCGUUCUUUCCUUGCACCA 3150
OY 2776 gaataatccaaagccaaatcgggggtgcatctgttgccttgctggaagaagaacca 2835
Db 3151 GACAUGGGCAAGGCGAAGAUUGCGCCCGCAUUCGACGACUUCGACGACGAAAGCCA 3210
OY 2836 aataatagacgagcagagcaagaagaagaagaagaagaagaagaagaagaagaagaaga 2895
Db 3211 CAAAUUGAUAAUUGGUCUAGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3270
OY 2896 ttctgagaagctctctctctctcatcatgctcgccagatgtttcatcctcgtgctta 2955
Db 3271 UUUAGGAACGUGCAUUCUAGAUACCGAGACCGCCGAGAAACAGCUCUUGCUGCGCGCUUG 3330
OY 2956 tccctcagatattgagcgaagaagaagaagaagaagaagaagaagaagaagaagaaga 3015
Db 3331 GACCCUGACCGUAAAGCUGACAAUUAUUGUGCGUUGCGACCCAGCGGUGUGGCAAG 3390
OY 3016 agcacttctgttcaactctctgagagactttagcccgctgcaaggaagaagtgtgttt 3075
Db 3391 AGUACCAACAUUGCAUUGGUCUUGAGCGCUUUUAGCAUUGCGCGGUCUACUUGU 3450
OY 3076 gatgtgtgtgagatgcaagaatgtagttagttagttagttagttagttagttagttagt 3135
Db 3451 GAUUGGAGAGCAUUAUAGUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 3510
OY 3136 cctcaagaagcctgtgtcttcaactgcaactgtgctggaagaacatgctatgtgacaac 3195
Db 3511 AGCCAGGAGCCGACACUGUACAGGAGGACAAUACAAGGAAAUUUAUUAUUAUUAU 3570
OY 3196 agccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3255
Db 3571 GAAAGUAGAGUACCGAGAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3630
OY 3256 ttatattgaagctctcctcgtgaagaatacaacaacaagaatgtagttagttagttagt 3315
Db 3631 UUCACAUUUGGUCUUGCGGAGGCGCUUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 3690
OY 3316 tctgctgagcagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 3375
Db 3691 UCUUGCGGCGCAAAAGCAACGUGGCAUUGGCGGAGCCUUCUUGGAGUCCCAAAUVC 3750
OY 3376 ttatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3435
Db 3751 CUUCUUCUUGCAUAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3810
OY 3436 gcccttgaataagcagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 3495
Db 3811 GCUUUGGAGUUGCGCGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 3870
OY 3496 attcagaacgagattttagttagttagttagttagttagttagttagttagttagttag 3555
```

```
Db 3871 AUUCAAAGCGCGAGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3930
OY 3556 catcaagagctcctgagaagaatcgagacatatatttaagttagttagttagttagttag 3609
Db 3931 CACAGCGAACUGGUCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3984
```

Search completed: April 22, 2002, 18:45:58  
Job time: 12941 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:32:15 ; Search time 9392.76 Seconds

(without alignments)  
142.602 Million cell updates/sec

Title: US-09-873-409-14

Perfect score: 3621

Sequence: 1 ttccgcttgcgtgactgact.....atgcacatcagtcagta 3621

Scoring table: IDENTITY\_NUC

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

EST:\*

1: em\_estfun:\*

2: em\_esthum:\*

3: em\_estlin:\*

4: em\_estom:\*

5: em\_estpl:\*

6: em\_estba:\*

7: em\_estro:\*

8: em\_estov:\*

9: em\_hic:\*

10: qb\_estl:\*

11: qb\_est2:\*

12: qb\_hic:\*

13: qb\_gss:\*

14: em\_gss\_fun:\*

15: em\_gss\_hum:\*

16: em\_gss\_inv:\*

17: em\_gss\_pln:\*

18: em\_gss\_pro:\*

19: em\_gss\_rtd:\*

20: em\_gss\_vit:\*

21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	545	15.1	545	10	AL040762 DKFP434C
2	386.6	10.7	943	10	AL520322 AL520322
3	385.8	10.7	405	10	AA243820 AA243820
4	378.8	10.5	872	12	AK020318 Mus muscu
5	357.4	9.9	559	11	BF692596 BF692596
6	298.8	8.3	894	11	BF584668 BF584668
7	275.2	7.6	726	11	BF629345 BF629345
8	273.2	7.5	1019	11	BG248052 BG248052
9	269	7.4	944	11	BF796582 BF796582
10	262.2	7.2	785	10	AV709991 AV709991
11	253.8	7.0	803	11	BG298756 BG298756
12	253.2	7.0	780	11	BG587938 BG587938

13	252.4	7.0	560	10	BE749379 BE749379
14	252	7.0	886	13	BH139685 BH139685
15	249.4	6.9	981	11	BF969667 BF969667
16	249.2	6.9	932	13	AZ670821 AZ670821
17	248.4	6.9	795	11	BF313560 BF313560
18	248.4	6.9	880	13	AZ687805 AZ687805
19	241.4	6.7	886	13	AZ540627 AZ540627
20	239.6	6.6	871	13	AZ682350 AZ682350
21	239.6	6.6	897	13	AZ541090 AZ541090
22	238.8	6.6	853	13	AZ679807 AZ679807
23	237.2	6.6	834	13	AZ548312 AZ548312
24	234	6.5	947	13	AZ683753 AZ683753
25	230.2	6.4	823	13	AZ532602 AZ532602
26	229.4	6.3	482	11	BI344244 BI344244
27	227.8	6.3	651	10	BE016688 BE016688
28	226	6.2	687	11	BG646725 BG646725
29	225	6.2	919	13	AZ690701 AZ690701
30	224.6	6.2	891	13	AZ682250 AZ682250
31	224.2	6.2	842	13	AZ671925 AZ671925
32	219.8	6.1	661	11	BG302675 BG302675
33	215.2	5.9	899	13	AZ348799 AZ348799
34	214.4	5.9	921	13	AZ687628 AZ687628
35	213.8	5.9	575	10	AA342644 AA342644
36	213	5.9	519	10	AA434959 AA434959
37	212.2	5.9	605	10	AV616675 AV616675
38	210.4	5.8	750	11	BG585786 BG585786
39	208	5.7	851	11	BF969062 BF969062
40	204.4	5.6	787	11	BG584063 BG584063
41	199	5.5	854	13	AZ682568 AZ682568
42	197.8	5.5	402	11	H82550 H82550
43	197	5.4	541	10	AI722858 AI722858
44	196.4	5.4	2676	12	AK014319 AK014319
45	195.6	5.4	704	10	BE239652 BE239652

## ALIGNMENTS

RESULT 1

AL040762 545 bp mRNA EST 29-FEB-2000

LOCUS DKFP434C1815 r1.434 (synonym: htes3) Homo sapiens cDNA clone

DEFINITION DKFP434C1815 5', mRNA sequence.

ACCESSION AL040762 GI:5409708

VERSION AL040762

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 545)

AUTHORS Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Bloeker, et al.)

JOURNAL Unpublished (1999)

COMMENT Contact: Bloeker H

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ). Email s.wiemann@dkfz-heidelberg.de.

Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

s1 sequence also available.

This clone (DKFP434C1815) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..545

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DKFP434C1815"

```

/clone_lib="434 (synonym: htes3)"
/issue_type="testis"
/development="adult"
/lab_host="DH10B"
/notes="Vector: pSPORT1; Site-1: NotI; Site-2: SalI"
BASE COUNT      171 a      89 c     138 g     147 t
ORIGIN

```

```

Query Match      15.1%; Score 545; DB 10; Length 545;
Best Local Similarity 100.0%; Pred. No. 6.9e-128;
Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 947 ccagatagaactcttccacagcttgatataaactgaatcatagaagaactgtg 1006
      |||
      1 ccagatagaactcttccacagcttgatataaactgaatcatagaagaactgtg 60
Qy 1007 aatttaaaatgttcttcaatataccatcaagaacatcatcaaatcttgaaggtc 1066
      |||
      61 aatttaaaatgttcttcaatataccatcaagaacatcatcaaatcttgaaggtc 120
Qy 1067 tgaatcccaaatataagtcggagagacagtcgcttgcgtcgaatgagcagtgga 1126
      |||
      121 tgaatcccaaatataagtcggagagacagtcgcttgcgtcgaatgagcagtgga 180
Qy 1127 agagtaagtaagtcgagctctgcagaggtatataatgcgagatgagcttcatag 1186
      |||
      181 agagtaagtaagtcgagctctgcagaggtatataatgcgagatgagcttcatag 240
Qy 1187 tgaatgagatgacatcaagacatttaaatgbcgacattatcgagacatattgagtg 1246
      |||
      241 tgaatgagatgacatcaagacatttaaatgbcgacattatcgagacatattgagtg 300
Qy 1247 ttgagtaagagcctgttcttgcggagccacatcaataatataagatgagagag 1306
      |||
      301 ttgagtaagagcctgttcttgcggagccacatcaataatataagatgagagag 360
Qy 1307 atgagtcgagatgagagagagagagagagagagagagagagagagagagagag 1366
      |||
      361 atgagtcgagatgagagagagagagagagagagagagagagagagagagagag 420
Qy 1367 tcatgagcttctcaataaattatcatattgtaggagaaagagagagagagagag 1426
      |||
      421 tcatgagcttctcaataaattatcatattgtaggagaaagagagagagagagag 480
Qy 1427 gaggagcgaagacagagagagagagagagagagagagagagagagagagagag 1486
      |||
      481 gaggagcgaagacagagagagagagagagagagagagagagagagagagagag 540
Qy 1487 ttctta 1491
      |||
      Db 541 tttta 545

```

```

RESULT 2
AL520322      943 bp      mRNA      EST      13-FEB-2001
LOCUS      AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CSDB006YCL5 5
DEFINITION      prime, mRNA sequence.
ACCESSION      AL520322
VERSION      AL520322.1 GI:12783815
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 943)
AUTHORS      Li, M.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France

```

```

FEATURES
source
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDB006YCL5"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/issue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact : Feng Liang life
Technologies, a division of invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

```

BASE COUNT      253 a      225 c     253 g     211 t
ORIGIN

```

```

Query Match      10.7%; Score 386.6; DB 10; Length 943;
Best Local Similarity 65.0%; Pred. No. 1.7e-87;
Matches 586; Conservative 0; Mismatches 314; Indels 1; Gaps 1;

```

```

Qy 2714 taatttactgcaattgcataatgagctatgagcactcgaaagacgtcgtttgctc 2773
      |||
      6 taatttactgcaattgcataatgagctatgagcactcgaaagacgtcgtttgctc 65
Qy 2774 ctgaaatccaaagacaaacgagcgagctgcagctcgtttgctgttgaaagaaac 2833
      |||
      66 ctgaaatccaaagacaaacgagcgagctgcagctcgtttgctgttgaaagaaac 125
Qy 2834 caaatatagacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2893
      |||
      126 ctgaaatccaaagacaaacgagcgagctgcagctcgtttgctgttgaaagaaac 185
Qy 2894 agttcgagaagctctcttcttccatcagtcgcccagatglttcatcctccgtgct 2953
      |||
      186 agttcgagaagctctcttcttccatcagtcgcccagatglttcatcctccgtgct 245
Qy 2954 tatccccaagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3013
      |||
      246 tatccccaagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 305
Qy 3014 aaagcattctgttcaactctgcagagacttatagaacccgagcagcagcagcagc 3073
      |||
      306 aaagcattctgttcaactctgcagagacttatagaacccgagcagcagcagcagc 365
Qy 3074 ttgattgtgtgagatgcaaaagatgaaatgagcagtggtccgttcccaatgcaatg 3133
      |||
      366 ttgattgtgtgagatgcaaaagatgaaatgagcagtggtccgttcccaatgcaatg 425
Qy 3134 ttccccaagagcctgtgtcttcaacccgagcagcagcagcagcagcagcagcagcagc 3193
      |||
      426 ttccccaagagcctgtgtcttcaacccgagcagcagcagcagcagcagcagcagcagc 485
Qy 3194 acagcgtgtgtgtgcatatagatgagatcaaaagacgagcagcagcagcagcagcagc 3253
      |||
      486 acagcgtgtgtgtgcatatagatgagatcaaaagacgagcagcagcagcagcagcagc 545
Qy 3254 ctttattgaagtcctccctgagagatacaacacagaagtgtgagcgtgaaagagcagc 3313
      |||
      546 ctttattgaagtcctccctgagagatacaacacagaagtgtgagcgtgaaagagcagc 605
Qy 3314 ttcttgagcgcagaaagacagcagcagcagcagcagcagcagcagcagcagcagcagc 3373
      |||
      606 ttcttgagcgcagaaagacagcagcagcagcagcagcagcagcagcagcagcagcagc 665
Qy 3374 ttctatgttgatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3433

```

```

Db      666 TTTTGGCTTTTGGATGAGCCACGCTGCTGATFACAGAAAGTAAAGGTTGGCCAG 725
Oy      3434 atgccttgataaagccaggacgggaagacatgccttagtggtaactcagagctctctg 3493
Db      726 AAGCCCTGACAAAGCAGAGAAAGCCGACCTGATGTATGTCACGCCGTGTCGA 785
Oy      3494 caattcgaagacgagattgtatgtg-gtctgcacaatggaagaagataaagaacaaga 3552
Db      786 CCATCCAGAAATGACAGACTTAATAGTGTGTTCAGAAATGACAGTCACAGAGATGCG 845
Oy      3553 acctcaagaagcctccagagaacacatatacttaagttagttgtaatcagatcagca 3612
Db      846 ACGCATCAGCAGCTGCTGGCAGACAAAGCATCTATTTTCAATGTCAGTGTCCAGGCT 905
Oy      3613 g 3613
Db      906 g 906

RESULT 3
AA243820 405 bp mRNA EST 06-AUG-1997
LOCUS z67906.r1 Soares.NbHMPu.S1 Homo sapiens cDNA clone IMAGE:668506 5'
DEFINITION similar to SW:MDRL_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;,
mRNA sequence.
ACCESSION AA243820
VERSION AA243820.1 GI:1874631
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS Hiller,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,D., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
,T., Waterston,R. and Wilson R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LML ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1572 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham.
FEATURES
Source Location/Qualifiers
1..405
/organism="Homo sapiens"
/db_xref="GDB:5562473"
/db_xref="taxon:9606"
/clone="IMAGE:668506"
/clone_id="Soares_NbHMPu.S1"
/lisue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site:1: Not I;
Site:2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHMP, pregnant uterus
NbHMPu, and fetal heart NbHMP19M) were mixed, and ss cicles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libaries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 115 a 81 c 85 g 124 t
ORIGIN

```

```

Query Match 10.7%; Score 385.8; DB:10; Length 405;
Best Local Similarity 99.5%; Pred. NO. 2.3e-87;
Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      822 tcttgctgttttctttagtgaatcacaagagttatgcatcttgagagagagctccta 881
Db      17 TCTCTGCTTTTCTTTTAACTGTAATCAGATGCAATTATTCATTTGAGACAGACGTCCTCA 76
Oy      882 ctctgaacacttcgcaatagcccgagagcgtgccttcatalattccaggttatgataa 941
Db      77 CTTTGAAACCTTCCGAATATAGCCCGAGAGCGCTGCTTCATATTTTCCAGGTTATGATATA 136
Oy      942 gaaaccagatataaactttccacagcctggatataaacttgatccatagaagaagac 1001
Db      137 GAACCCAGATATAGATCTTTCCACAGCGTGGAATTAACCTGAATCATAGAAGAGAAC 196
Oy      1002 tctggaatttaaaaatgtctcttccttcattatccatcaagaaccatctcaagattcga 1061
Db      197 TGTGGAATTTTAAATATGTTTCTTTCAATTATTCATCAAGACCATCTATCAAGATTCTGAA 256
Oy      1062 aggtctgaatcctcagaatagctctgagagacagctgcgcttgctgggtlctcaatggcag 1121
Db      257 AGGTCTGAATCTCAGAAATTAAGTCTGAGAGACAGTCCGCTTGTCGCTCAATGCGCAG 316
Oy      1122 tgggaagagtaacggttagtcacagctctcagagaggttatatgatccggatgatgcttca 1181
Db      317 TGGGAAGAGTACGGTAGTCACACCTTCTCAGAGGTTATATGATCCGATGATGAGCTTAT 376
Oy      1182 catggtgatgagaatgagatcagagcct 1210
Db      377 CATGCTGATGAGATGACATCAGAGCTT 405

RESULT 4
AK020318 872 bp mRNA HTC 05-JUN-2001
LOCUS AK020318
DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
library, clone:9230106F14, full insert sequence.
ACCESSION AK020318
VERSION AK020318.1 GI:12860872
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
clone_lib:RIKEN Full-length enriched mouse cDNA library
clone:9230106F14.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 872)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
METHODS Methods in enzymology. 303, 19-44 (1999)
JOURNAL 992799253
MEDLINE 10349636
PUBMED 10349636
REFERENCE 2 (bases 1 to 872)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 20499374
REFERENCE 3 (bases 1 to 872)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,T., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
Sunii,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasaiyagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipillary sequencer

```



```
se_1: NOT1;
```

2.

-cata 922

982 982

Magac 1042

2011 7776-  
111

—  
—  
—

111

111  
5ATCG 457

PGTCA 517

11/15/2005 3:11

Db	578	GTGAGAGAGGGGGCGAGCTGAGTGGGGGACAGAAACAGAGATCCG-ATTGCCGGGGGCC	636
OY	1463	tagttcgaaccaccaagaltctgatttagatgaggtactgctctgcctgattcaaga	1522
Db	637	TGCTCCGCAATCCCAAG-TGCCTTTCGTGGAGCAAGCAACCTCCAC-CTGTGATACAGAAA	694
OY	1523	gcaagtgcctcttcaagctgacatgcgaggaagcgagcaagaagtgcg	1569
Db	695	GTGAAGCTGGGGTTCAGCGCCGACCTAGATAGGATACAAAGGCGNG	741

  

RESULT	7
LOCUS	BG293345
DEFINITION	602390738F1 NIH_MGC_94 Mus musculus CDNA clone IMAGE:4502552 5', mRNA sequence.
ACCESSION	BG293345
VERSION	BG293345.1 GI:13052943
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cga@b-remail.nih.gov">cga@b-remail.nih.gov</a> Tissue Procurement: The Cepko Laboratory CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLNL0371 row: 1 column: 09 High quality sequence stop: 658.

  

FEATURES	source
Location/Qualifiers	1..726
/organism="Mus musculus"	
/db_xref="taxon:10090"	
/clone="IMAGE:4502552"	
/clone_1lb="NIH_MGC_94"	
/l1ssue_type="retina"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: eye; Vector: pCMV-SPOrt6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
BASE COUNT	189 a 220 c 190 g 127 t
ORIGIN	

  

Query Match	7.6%; Score 275.2; DB 11; Length 726;	
Best Local Similarity	68.3%; Pred. No. 3.8e-59;	
Matches	396; Conservative 0; Mismatches 183; Indels 1; Gaps 1	
OY	2984 tagcattgtggggagcagcggctgttgggaaagcaactctgttcaactctgcagagac	3043
Db	3 TGGCCCTGGTGGCGACGATGGCTCGGGAMAAGCACATGGTCCAGCGTCGAGCGCT	62
OY	3044 ttatgaccgccgtgaaagcaagatgcctgttggatggtgtgtatgcaagaatgaaat	3103
Db	63 TCTATGACCCCATGCTGTGATGAGTCTTTAGATGAGTCAAGACGAMAAGAACTCAATG	122
OY	3104 tacaagtgcctcgttccccaatagacaatcgttctccaaagagcctgtgctctcaactga	3163
Db	123 TCCAGTGGGTCGAGCTCAACTGGGCAATTTGTCTCCAGSAAACCAATTCCTTGAAGTGA	182
OY	3164 gcaatgtcgagaacatcgactatggtgtgacaacagccgtgtgtgtccatagatgatga	3223

Query Match	7.5%	Score 273.2	DB 11	Length 1019
Db	183	GCATGCGAGAGAACTCCGCTATTGGAGACAAAGCGCGGTCTGGCTCATGTATGAGATTG	242	
Oy	3224	aagaagccgcaaatgacgaataatccattcttatttgaagggtccctgagaataata	3283	
Db	243	TGAGGCGAGCCAAAGGAGGCGCAACATCCACCCCTTCATCGAGACGCTGGCCCCAAAATAATA	302	
Oy	3284	acacgaagtgtgactgtaaggagacagcttctcggcgccagagaaacaaactagcta	3343	
Db	303	ACACAAGAGTATGAGACAAAGGGGAGCGACACTCTGTGGGGGCCAGAGACGAGATTGGCA	362	
Oy	3344	ttgcaagggtctcttccaaaaccacaatttatttgttgatgtaggcattcagccc	3403	
Db	363	TGCGCCGAGCCCTCATTCAGACAGCGCTCGGGTCTCTACTGCTGATGAAGCCAGCTAGCTC	422	
Oy	3404	tgcataatgacacgtagaagggtggttcagcatgcgccttataaagccagacggaagga	3463	
Db	423	TGGATCTCTAGAGTGAAGAAAGTTGTTCCAGGAGACACTGGACAAAGCCAGGAGGCGCCA	482	
Oy	3464	catgcctggtgttactactcaaggtctctcgaatctcagaacgcag attgtagtgggt	3522	
Db	483	CCTGCATTGCGATTCCTCACCGCTGTCACCATCATCAGAACCGGCACTGTATTGTGTG	542	
Oy	3523	ctgcacaatgtaagaataaagaacaagaactcacaag 3562		
Db	543	ATTGAGAACGCGAAGTCAAGSAGCAGCGCACCCACCCAG 582		
RESULT	8			
LOCUS	BG248052			
DEFINITION	603235987F1 NCI_CGAP_Mam1 Mus musculus CDNA clone IMAGE:4488404	5'		
ACCESSION	BG248052			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 1019)			
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cga@bbs-rcmail.nih.gov			
	Tissue Procurement: Gilbert Smith, Ph.D.			
	cDNA Library Preparation: Life Technologies, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: L1AM10334 row: n column: 21			
	High quality sequence stop: 650.			
FEATURES	Location/Qualifiers			
source	1..1019			
	/organism="Mus musculus"			
	/strain="FVB/N"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:4488404"			
	/clone_lib="NCI_CGAP_Mam1"			
	/tissue_type="tumor, biopsy sample"			
	/dev_stage="3 months, virgin"			
	/lab_host="DH10B"			
	/note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;			
	Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.			
	library constructed by Life Technologies. Investigator			
	providing samples: Gilbert Smith, NIH"			
BASE COUNT	276 a 262 c 283 g 196 t 2 others			
ORIGIN				

Best Local Similarity 63.2%; Pred. No. 1.3e-58;  
Matches 437; Conservative 0; Mismatches 253; Indels 2; Gaps 1;

```

OY 2881 gaaggaattgaggttcgagaagctctctctctacatctgcgcagatgttcc 2940
Db 4 GAAGGAAAATGTCATTTAGTGGAGTCGTCACTATCCACCCGACCCGACATCCCA 63
OY 2941 acccccggtggttacccctcagatctgagcgaggaagaacagtagacttctggagac 3000
Db 64 GTCCTTCAAGGGGCTGAGCCTTGAAGTGAAGAGGCCACACCTGGCCCTGGTGGCAGC 123
OY 3001 agcggctgtygggaaagcaactctgttcaactctgcagagactatgaccgcgtgcaa 3060
Db 124 ACTGCTGCGGGAAGAGACAGTGTGTCAGTGTGAGCGGCTTACAGACCCCATGGCT 183
OY 3061 ggcagagtgcttctgagtggtgatagcacaagaatgaatgatagtgtgcgtcc 3120
Db 184 GGATCAGTGTCTAGTATGGCAAGAAATGAAGCAACTGAATGTCAGTGTCCGAGCA 243
OY 3121 caataagcaatcgtccctcaagaagcgtgtcttcaactctgcagactctgagaacatc 3180
Db 244 CAGCTGGGATGTGTGTCACAGGCCCATCTCTTTGACTGACGATCCAGAGAACTT 303
OY 3181 gctatgtgtgacaagccgtgtgtgcatagatgagatcaagaagccgcgaatgca 3240
Db 304 CCTGACGAGACACAGCCGGGCTGCTGTATGAGGATTTGTGAGGACCAAGAG 363
OY 3241 ggaatattcatctcttcttctgaaagctcctctgagaataacacacaagtgtgacg 3300
Db 364 GCCAATCTCCACCTGCTGCGGAGGAGCAAGCAAGCATGATTAACACAGAGTAGAGAC 423
OY 3301 aaagagacagactcttctgagcgccagaaacagactagatctgcaaggctctctc 3360
Db 424 AAAGCAGCTACCTGTGCGGAGGAGCAAGCAAGCATGATTAACACAGAGTAGAGAC 483
OY 3361 caaaaccccaaatcttcttctgagtgagcgcaactcagccctcgataatgacagtga 3420
Db 484 ACACAGCCTCACATTTTACTTCTGAGCAAGCAACATCAGCTGTGATACAGAACTGA 543
OY 3421 aagtggtctgacagatgcctgtgataaagccaggaacggaagacatgctagtgtact 3480
Db 544 AAGGTTGTCAGGAAGCGGTGAGCAAAAGCCAGGAGGCCGACCTGATTCGATCGCT 603
OY 3481 caccagctctctgcaattcagaagcagatctgatagtgtt--cgcacaatgagaaga 3538
Db 604 CACCGCTGTCCACCATCCAGAAACGGGACTGTGATGCGTGTGATTCAGAAACGGAGG 663
OY 3539 taaaggaacaggaactcatcagaagctcgt 3570
Db 664 TCAGGGAGCAGCGGACCCCAACAGCAAGTG 695

```

## RESULT 9

```

BF796582      944 bp      mRNA      EST      12-JAN-2001
LOCUS         602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
DEFINITION   mRNA sequence.
ACCESSION    BF796582
VERSION      BF796582.1 GI:12101636
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 944)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
              Email: cgaabs-remail.nih.gov
              Tissue Procurement: Louis Staudt, M.D., Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLM955 row: n column: 15  
High quality sequence start: 9  
High quality sequence stop: 669.  
Location/Qualifiers

FEATURES  
source 1..944

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4341710"  
/clone\_1b="NIH\_MGC\_85"  
/tissue\_type="Lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lymph; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.867 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 265 a 241 c 237 g 201 t  
ORIGIN

Query Match 7.4%; Score 269; DB 11; Length 944;  
Best Local Similarity 65.4%; Pred. No. 1.5e-57;  
Matches 409; Conservative 0; Mismatches 215; Indels 1; Gaps 1;

```

OY 2779 tattcaaaagccaaatcggggctgcgcacatctgttgccttctgtaaaagaacccaat 2838
Db 39 TATGCTAAAGCTTAAGCTGCTCGAGCCCACTTATCATGCTGTTTGAAGACAACTCTG 98
OY 2839 atagaagccgagtaagaaggaagaaagccagacacatgtaaggaatttaggttt 2898
Db 99 ATTGACAGCTACAGTGAAGGGGCTGAAGCCGTGATTAATTGAAGAAATATACATTT 158
OY 2899 cgagaagctctctctctctatccatctgcgccagatgttctcaactcgtgtctatcc 2958
Db 159 AATGAAGTGTGTTCAACTATCCACCCGAGCAAAAGTGCAGTGTCTCAAGGGCTGAGC 218
OY 2959 ctcaagattgagcgaggaagacagtagaatgttgggagcagcgctgtggaagaaagc 3018
Db 219 CTGAGAGTGAAGAAAGCCAGACACATACCCCTGTGTGGGAGCAGTGTGTGGAAGAGC 278
OY 3019 acttcgtcaactcttctgcagagacttatgaccccggtgcaaggaagaatgctgttga 3078
Db 279 ACGGTGTCCAGCTCCTGAGCGGCTCTACGACCCCTGTGGGGGACAGTGTTCGAT 338
OY 3079 ggtgtgatagcacaagaatgtaatgatacagtgagctgcgttcccaatagcaatcgttct 3138
Db 339 GGTCAAGAGCAAAAGAACTCAATGTCCAGTGTGCTCAGAGCTCAACTCGGAATCGTGT 398
OY 3139 caagagcctgtgctcttcaactgacagcatgtctgagaacatgcgctatgtaaacagc 3198
Db 399 CAGGACCTATCTCTATTTCAGTCAAGCATTTGCCGAATAATTCCATATGAGAGCAACAGC 458
OY 3199 cgtgtgtgcatatgatagatacagaagaagccgcgaatgagcaaatctcattttt 3258
Db 459 CGGTTGTATCCAGGATGAATCTGTGATGTCAGCCAAAGCTCCCAACATACATCTTTC 518
OY 3259 atgtaagctctccctgagaataatacaacacacaagcttgagctgtaaaagagacagcttct 3318
Db 519 ATGGAACGTTACCCCAAAATATGAACAAGTGGAGATTAAGGGAGCTCAGCTCTCA 578
OY 3319 ggcggccagaacaaagactagctatgcaaggctctcttccaaaacc-aaatttt 3377
Db 579 GGAAGTCAACAACAGAGGATTTGCTATTTGCCCGAGCCCTTCAGACAACTCAAAATCT 638
OY 3378 atgtgtgtagggccacttagcc 3402
Db 639 CCTGTTGATGAAGCTACATCAGCC 663

```



RESULT	10
LOCUS	AV709991
DEFINITION	AV709991 Cu Homo sapiens cDNA clone CUAJA09 5', mRNA sequence.
VERSION	AV709991
KEYWORDS	EST.
SOURCE	human
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 785)
REFERENCE	Peng Y., Song H., Gu Y., Yang Y., Gao G., Xiao H., Xu X., Li N., Qian B., Liu F., Qu J., Qu J., Gao X., Cheng Z., Xu Z., Zeng L., Xu S., Gu W., Tu Y., Jia J., Fu G., Ren S., Zhong M., Lu G., Hu R., Chen J., Chen Z. and Han Z.
AUTHORS	Homo sapiens cDNA Cu clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel.: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.
TITLE	JOURNAL
COMMENT	Location/Qualifiers 1..785 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="CUAJA09" /clone_1lb="Cu" /tissue_type="adrenal cortico adenoma for Cushing's syndrome" /dev_stage="Adult" /note="Vector: pBluescript sk(-)"
FEATURES	source
BASE COUNT	210 a 180 c 211 g 183 t 1 others
ORIGIN	
Query Match	7.2%; Score 262.2; DB 10; Length 785;
Best Local Similarity	64.0%; Pred. No. 7.9e-56;
Matches 458; Conservative	0; Mismatches 254; Indels 4; Gaps 4
Dn	2704 ggcatgtcatagatgttttaactcgoaatlgtcatatggagctagtgcacatcggaanaacgctc 2763
Dn	73 GGAAATGGAGAAAGTATTTCACGCTGTGTTTGTCGCCATGCGCCGTGGGGCAAGTCAGT 132
Oy	2764 gtlttgccctccgaatatcccaaggccaatcggggcgctgcgatctgtttgcctgttg 2823
Dn	133 TCATTTGCTCCTGCACATCATGACCACAAGCCAATAATATCACGACCACATCATGATGATCATT 192
Oy	2824 gaagaagaaccacaatatagacagccgcagtcacagaagaaggaagaagcacagacacattgaa 2883
Dn	193 GAAAAAACCCCTTTGATTGACACGCTACACAGCAGGAAGGCTTAATGCCGAACACATTTGGA 252
Oy	2884 gggaaatltagagtlcttgagagaagctcttcttcattcatcatalgtgccagagatgtttcatac 2943
Dn	253 GGAATGTCACTTTGGTGAAGTTGTAATTCACACTATTCACCCAGCCGACCGACATCCAGT 312
Oy	2944 ctccgttgtctalcacctcagtaltagcgagagaagaacagtagcaattgtgggagcagc 3003
Dn	313 CTTCAGGAGACTGAGCTGGAGTGGAAGAAGGGCCACAGCGCTGCTGGTGGGACACAGT 372
Oy	3004 ggcgtggggaaaagaccctctgttcaactctctgcagagacttatagaccgcgtgcaagaa 3063
Dn	373 GGGTGTGGGAAGACACAGTGTCTCAGCTCTCTGAGGGGTTCTTACGACCCCTTGGCAGGG 432
Oy	3064 caagtgctgttattgattgtgtgtagtgcnaaagaatattaatlaagttgctcgcgltcccaca 3123
Dn	433 AAAGTGTCTGTTATGGCACAAGAAATTAACGACGATGAATTTTCAGTGGCTCCGAGCAC 492
Oy	3124 atagcaatcgtlltcccaagaagcctgtgtcttcaactgtcagactgcagcaatltgtcgaacaatgcgc 3183

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	493	CTGGGCGATCGTGTCCAGAGAGCCATCTCTGTTTACTGACGACATGTGTGAGAACATTGCC	552					
Qy	3184	tatgtgtcaacaacgccc-gtgtgtgtgcatatagatgatcaaaagacgcgaatgtcagc	3242					
Db	553	TATGAGAGCAACACACGCGGGGTGTGTGTACAGAGAGAGATGTGTGAGGCGACCAAAAGGAGGC	612					
Qy	3243	aaatatcatctcttattatgttagagctccctcgagaataaacaacacaagttgtgactgaa	3302					
Db	613	CAACATATACATCCCTTTCATCGAGTCAGTCGTCCCT-ATTAATATATGACACTAATACAGACAA	671					
Qy	3303	aggagcacagctctctgtgcgcgcagaacaagactagctatctcaaggcctctctcca	3362					
Db	672	AGGACCTTAGCTCTCTGTGGGGGCCAAGAAACACA-TTGCCATAGCTTGTCGCCCTGTGAGA	730					
Qy	3363	aaaaccccaaatattatgtgtgatgagagccacttcagccctgataatgacagt	3418					
Db	731	CAGACTTATTA-TTTGTTTGTGATGAGACGACGCTCAGCTCTGTATGACAAAGTG	785					
RESULT	11							
LOCUS	BG298756	803 bp	mRNA	EST	21-FEB-2001			
DEFINITION	602396681p1 NIH_MGC_94	Mus musculus	cdna	clone	IMAGE:451157 5'			
ACCESSION	BG298756							
VERSION	BG298756.1	GI:13063728						
KEYWORDS	EST.							
SOURCE	house mouse.							
ORGANISM	Mus musculus							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.							
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)							
COMMENT	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: cgabs@emal.nih.gov							
	Tissue Procurement: The Cepko Laboratory							
	cDNA Library Preparation: Life Technologies, Inc.							
	DNA Sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LNL at:							
	http://image.lnl.gov							
	Plate: L1AM10394 row: b column: 22							
	High quality sequence stop: 698.							
FEATURES	Location/Qualifiers							
source	1..803							
	/organism="Mus musculus"							
	/db_xref="taxon:10090"							
	/clone_image="IMAGE:451157"							
	/clone_lib="NIH_MGC_94"							
	/tissue_type="retina"							
	/lab_host="DH10B (phage-resistant)"							
	/note="organ: eye; Vector: pCMV-SPORT6; Site:1: NotI;							
	Site:2: SalI; Cloned unidirectionally; oligo-dT primed.							
	Average insert size 3.3 kb. Library enriched for							
	full-length clones and constructed by Life Technologies.							
	Note: this is a NIH-MGC Library."							
BASE COUNT	240 a 162 c 205 g 196 t							
ORIGIN								
Query Match	7.0%; Score 253.8; DB 11; Length 803;							
Best Local Similarity	60.9%; Pred. No. 1.le-53;							
Matches	485; Conservative 0; Mismatches 302; Indels 10; Gaps							
Qy	535	tcatgaccagtaagaatgaatgaatccatattccaaagctgggctggtgacagaagaatgc	594					
Db	15	TCATTACTGATGAAGCAACTCCATGCTTATGCAAAAGCTGGAGCACTTCTGAAGAAGTC	74					
Qy	595	tgtctatcaatccgaacagtcataagctttagagggccagagaagaactccaagtat	654					

```

Db      75 TTAGAGGCAATCAGAACTGTGATTTG-CGTTGGAGAGCAAAAGAAAGAACTTGAAGGATAC 133
Oy      655 acacagaaatcctcaagaatgcaagaagatttggcataaagaagcattagcttcaaaagt 714
Db      134 AATAACCACTTGGGAAGAGCTTAAAGGCTGGGGATTAAGAAAGCTATACCGGCCAACATTC 193
Oy      715 tctctgtgtcgtgtactcttcttgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 774
Db      194 TCACAGGGGTGACGCTTTCTCTTATCTATGATCATATATGCTTGGCATTTCTGGTATGGG 253
Oy      775 aacctctgattcttaatgaagaaccttgatataacacaggaactgtcttctgtcttcttcttct 834
Db      254 ACTTCTGTGATCTCCAAAGAA-----TACTCTATGAGCAAGTGTCTACGTCTCTTC 306
Oy      835 tttagtgaatccatagaagaatgattatgcttggagcagcagcagcagcagcagcagcagcagc 894
Db      307 TTTTCCGTTAATTTGGAGCATTCAGTGTGGACAGCATCTCCAAATATTTGAACCTTC 366
Oy      895 gcaatagcccgagagcgtgctcttcatatattccaggtatattgaagaaccagatata 954
Db      367 GCCAATGCACAGAGAGGAGCTTATGATAGCTTCAAAATATATTGATTAAGCCCAATATA 426
Oy      955 gataactttccacagcgtgataataaacctgaatccatagaagaagcagcagcagcagcagcagc 1014
Db      427 GACAGCTTCTCAAAAGAGTGGGCAACAAACAGACATACAAAGAAATCTGGAATTTAAG 486
Oy      1015 aatgtctttcattcaatcattcaacgaaccatcattcaagaatctgaaagcttgaatctc 1074
Db      487 AATATTCCTCAGTTCAGTTCACCTCGAAAGAAAGTTCAGATCTTGAAGGCGCTCAATCTG 546
Oy      1075 agaattaagcttcgagagcagcgtgctgcttgcgttccatagcagcagcagcagcagcagc 1134
Db      547 AAGGTGAAGAGGAGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
Oy      1135 gtagcagcttcctgagaggttataatgatcgcagatgagtgcttattatgagtgagtgag 1194
Db      607 ACTGTCAGAGTGAATCAAAAGGCTGACAGCCCTTA-GATGGCATGCTGATGCTGATGCTGATG 665
Oy      1195 aatgaatcagaagccttaaatgctgcgcattacgcagacacatattgagtgagtgagtgag 1254
Db      666 CAGGACATCAGAACCATCAATGATGAGGATGATGAGGAGGATGAGGAGGATGAGGAGGATGAG 725
Oy      1255 -gagccttcttctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1313
Db      726 GGAACCTGCTTTGTTTGCACACCGAGATCGGAGAAATTCGTTATGAGGCGAGAGAGT 785
Oy      1314 gactgataagaagatg 1330
Db      786 GGTTCCTGCTGATGATG 802

```

RESULT 12  
Bg587938  
LOCUS Bg587938 780 bp mRNA EST 11-APR-2001  
DEFINITION Bg587938 MHAM Medicago truncatula/Gliomus versiforme mixed EST  
ACCESSION Bg587938  
VERSION Bg587938.1 GI:13603002  
KEYWORDS  
SOURCE Medicago truncatula/Gliomus versiforme mixed EST library.  
ORGANISM Medicago truncatula/Gliomus versiforme mixed EST library.  
REFERENCE 1 (bases 1 to 780)  
AUTHORS Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.  
and Fraser,C.M.  
TITLE ESTs from roots of Medicago truncatula after colonization with  
Gliomus versiforme, 2001  
JOURNAL Unpublished (2001)  
COMMENT Contact: Harrison M.J.  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73401

```

FEATURES
Source
    source
        /organism="Medicago truncatula/Gliomus versiforme mixed EST
        library"
        /cultivar="Medicago truncatula genotype A17"
        /db_xref="taxon:119092"
        /clone="PMHAM-51L19"
        /clone.lib="MHAM"
        /issue.type="roots colonized with Gliomus versiforme"
        /dev.stage="Roots harvested at 10, 17, 22, 31 and 38 days
        post-inoculation with Gliomus versiforme. The library was
        made from a mixture of RNA from each of these stages."
        /lab.host="E. coli strain XL0R"
        /note="Vector: p Bluescript SK-. Site.1: EcoRI; Site.2:
        XhoI; CDNA was prepared from polyA+ enriched RNA from
        roots harvested at 10, 17, 22, 31 and 38 days
        post-inoculation with Gliomus versiforme. The CDNA was
        directionally ligated into the UniZap XR vector from
        stragene and packaged using Gigapack III Gold packaging
        extracts. Plasmids containing CDNA inserts were excised
        from the recombinant lambda-Zap phage using Ex-assist
        helper phage and propagated in XL0R cells."
BASE COUNT 234 a 156 c 184 g 206 t
ORIGIN
Query Match 7.0%; Score 253.2; DB 11; Length 780;
Best Local Similarity 58.5%; Pred. No. 1.6e-53;
Matches 441; Conservative 0; Mismatches 313; Indels 0; Gaps 0;
Oy      863 ttgagcagcagctccctcaacttgaaccttcgaatgagccgagagagcttctcata 922
Db      4 TGGGCAATCTGCACGACGATGCTGCATTTACAAAGCTAGAGTGGGCTGCTGAAGA 63
Oy      923 ttctcaggtatgataagaaccagtatagataaactttccacagcttgatataac 982
Db      64 TTTTCGGATTAATGATATACCAAGCTGATGATGATGATGATGATGATGATGATGATGAT 123
Oy      983 ctgaatcattagaagaagccttgagatcttaaaatgcttcttcaatcattcaagaac 1042
Db      124 TAGAGACAGTTACTGACTTGTGAACTGMAAAAGTGAGCTGCTTCTTATCATCAAGAC 183
Oy      1043 catctataagaagcttgaaagcttgaaatctcagaatcagaatgaagcttgagagagcagcct 1102
Db      184 CTGAAGTTCTGATCTCAATGATTTCTCTTGTAGTGTCTCTGCGGAAAGACATAGCTT 243
Oy      1103 tgcgtgctcaatgagcagcttgagagagtagagtagtcagcagcttctgcagaggtatata 1162
Db      244 TAGTTGTAAGACAGGCGCTCTGCAAGAGACAGCTGTTCTTCTTATGAGCAATTTATG 303
Oy      1163 atccgagatgagcttcatcagtgtgagatgagaatgaatcagacagccttcaatgtgcgc 1222
Db      304 ATCCAATCTCAGACAGAAATGTTGATGGGATGAGCATGATTAATCACTTGAAGCTTAAT 363
Oy      1223 attacgagacatatgtgagtggttagtcaagaagcctgttctgctggagacacatca 1282
Db      364 GGTGAGGCAACAAATGAGACTAGAGCCCAAGAACTGCTTTGTTTGGCCACAGATTC 423
Oy      1283 gtaacaatacaagatgagcagagatgctgaacgagatgaagaagaatgagagcagca 1342
Db      424 GAGAAATATATCTCTTGGGAAGCCTGATGCAAAACAGAGTTGAGTGAAGAGCTGCTA 483
Oy      1343 gggaagcaaatgctgataattatcattgagagtttccctaataaattatcatattgag 1402
Db      484 GGGTTGCTAAATGCTCATTCATCATCAATCAAACTTCGAAAGGCTTTTGAACCTCAGTAG 543
Oy      1403 gggaagaagagctcaaatgagtgagagggcagaacaagaagatgcgaattctgtgct 1462

```

```

DB 544 GAGAAAGAGACTACTATTTCTGGAGGACAAAACAAAGAAATAGCAATAGCAAGGCAA 603
1463 tagtctgaaccccaagatctgatttagatgaagctgaagctgcccctgattgagaa 1522
DB 604 TCGTAAATATCCAGCAATTTCTCTCTCGATGAGCAACAACTGCAATGACTCTGAAT 663
1523 gcaatcagctgtctcaagctcagctgagaagcgagcaaaagctggaactacaatcgtag 1582
DB 664 CAGAAAAGCTGCTGACAGAAAGCACTTGACCGGTTCATGATTTGACAGACAACTTGTA 723
1583 tagcacacgactcttctactactgaagtagcaga 1616
DB 724 TTGCTCATGCGCTCTCCACAAATTCGCAAAAGCTGA 757

RESULT 13
BE749379 560 bp mRNA EST 25-APR-2001
LOCUS BE749379 200194 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BE749379
ACCESSION BE749379
VERSION BE749379.1 GI:10163371
KEYWORDS EST.
SOURCE
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
1 (bases 1 to 560)
Smith,T.P., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrnenkrug,S.C., Bennett,
G.T., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perteaux,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
Keeler,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR PRIMERS
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 108 row: F column: 11
Seq primer: ATTTAGGTGACACATATG.
FEATURES
Source
1..560
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 4BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 20 and day 40
embryos."
BASE COUNT 154 a 153 c 146 g 107 t
ORIGIN
Query Match 7.0%; Score 252.4; DB 10; Length 560;
Best Local Similarity 67.5%; Pred. No. 2.3e-53;
Matches 355; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
OY 3090 aaagaattgaatgacagtggtcgtccgtcccaataatagcaatcgtctcagaagccgtg 3149
DB 1 AAAGAGCTCAACGTCAGTGCTCAGAGCCCACTTGGAATCGTTCGACGAGACCCCGT 60

```

```

OY 3150 gctctcaactcagcagcatgctgtagaagatcgctctatgtagaagaagccgtgtgccc 3209
DB 61 CCGTTTGAAGCTGAGCATTTGCGACCAACAACTTGTGGGACAAACAGCGGCTGTAC 120
OY 3210 attagatgagataaagaagccgcaatgagcaaatatccattctttagaagctc 3269
DB 121 CATGCCCTGAATTTGTGAGCGGACGCAAGCGCCAAATCATCTTTCATTAGACTT 180
OY 3270 cccttgagaataacacacacaagaattgactgaaaggagcagcttctgycgagcaaa 3329
DB 181 GCCCCCAAAATATGAACAAAGAGTGGAGATTAAGGGGACTCAGCTTCCGGGGACAGAA 240
OY 3330 acaagaactgctatgcaagggctctctccaaaaaaccataattatgttgatga 3389
DB 241 ACAGAGGATTTGCTATTGCCCCGAGCCCTTCATCCGACACCCCGGATCTTACTGTGATGA 300
OY 3390 ggcgcactcagccctcgatataatgacatggaaggtgtgttcacagatgccttataaagc 3449
DB 301 AGCCACGTCAGCACTGATCTGAAGTGAAGATGATGTCACAGAGACCCCTGGACAAAGC 360
OY 3450 cagagcgggaagagacatgctcagtagtgcactcacaggtctctgcaattcagaagcaga 3509
DB 361 CCGAGAGCGCGCGACCGCTGCAATTCGATGCTCAACCGGCTGTCACCAATCCAGAACGAGA 420
OY 3510 ttgatatgtgtgtctgcacatggaagaataaaggaaacgaactatcataagagctcct 3569
DB 421 CTGTATGATGTTGATTTGAGATGCGACAGATCAGGAGCACGGCACACACAGAGCTGCT 480
OY 3570 gagaatcagagacatataatttaagtagtagatggaatgcacagtcagt 3615
DB 481 GGCAACAGAAAGGCAATCTATTTCACCATGTCAGTGTCCAGGCTGG 526

```

```

RESULT 14
BH139685 886 bp DNA GSS 07-AUG-2001
LOCUS BH139685
DEFINITION BH139685 Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
ACCESSION BH139685
VERSION BH139685.1 GI:15098746
KEYWORDS GSS.
ORGANISM Entamoeba histolytica.
Entamoeba histolytica.
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 886)
Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
TITLE HMI:IMSS sheared DNA library (2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: entae@ligr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 18
High quality sequence stop: 851.
FEATURES
Source
1..886
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHS01; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 23:07:44 ; Search time 866.72 Seconds  
(without alignments) 3500.964 Million cell updates/sec

Title: US-09-873-409-14

Perfect score: 3621  
Sequence: 1 ttccgcttgcgtatgact.....atgcacagtcagtcagctga 3621

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N.Geneseq\_1101.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1325.4	36.6	4279	22	AAD03504
2	1325.4	36.6	4279	22	AAD03505
3	1325.4	36.6	4279	22	AAD03506
4	1323.8	36.6	4279	22	AAD03488
5	1320.6	36.5	4317	22	AAD03489
6	1316.2	36.3	4349	22	AAH57442
7	1316.2	36.3	4646	15	AA072872
8	1316.2	36.3	4646	21	AA294738
9	1314.8	36.3	6505	17	AA13394
10	1314.8	36.3	4669	14	AA052726
11	1314.6	36.3	3860	21	AAZ49332

12	1314.6	36.3	4669	8	AAH07052
13	1314.6	36.3	8630	21	AAZ24042
14	1313	36.3	3860	21	AAZ49333
15	1311.4	36.2	8630	21	AAZ24041
16	1309.8	36.2	4378	11	AAQ04522
17	1308.2	36.1	4669	19	AAV32645
18	1306.6	36.1	4186	22	AAH6127
19	1306.6	36.1	4195	22	AAH6128
20	1305.2	36.0	4264	19	AAH6533
21	1303.6	36.0	4264	19	AAH6534
22	1300.2	35.9	3988	21	AAZ88973
23	1276.2	35.2	3924	21	AAZ94742
24	1276.2	35.2	3924	21	AAZ88974
25	1216.8	33.6	4425	21	AAZ53048
26	1216.4	33.6	4189	21	AAZ49334
27	1216.4	33.6	4313	14	AAQ38950
28	1215.2	33.6	4369	21	AAZ52047
29	1212	33.5	4233	21	AAZ90198
30	1212	33.5	4233	22	AAE27498
31	1204.2	33.3	4788	21	AAZ49335
32	863.4	23.8	4776	21	AAZ94744
33	856	23.6	2726	15	AAQ70907
34	856	23.6	2726	15	AAQ70916
35	856	23.6	2726	18	AAZ43322
36	598.8	16.5	4175	20	AAV69392
37	511.4	14.1	4102	21	AAZ45942
38	496.8	13.7	4002	20	AAH85823
39	496.4	13.7	4051	21	AAH09019
40	437.4	12.1	4047	20	AAH60201
41	392.2	10.8	4224	19	AAV42347
42	383.8	10.6	2698	20	AAV69394
43	365.4	9.8	4800	20	AAH60202
44	355.8	9.8	3512	20	AAV69395
45	352.6	9.7	3923	17	AAZ35230

## ALIGNMENTS

RESULT 1	
AA03504	
ID AAD03504 standard; cDNA; 4279 BP.	
XX	
AC AAD03504;	
XX	
DT 13-JUN-2001 (first entry)	
XX	
DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.	
XX	
KM Dog: P-glycoprotein allelic variant; multidrug transporter; MDR1;	
KW drug bioavailability; transgenic animal; genetic model; ss.	
OS Canis familiaris.	
XX	
OS	
XX	
FT Key	Location/Qualifiers
FT CDS	17..3862
FT	/*tag= a
FT	/product= "Dog P-glycoprotein (PGP) allelic variant
FT	(Genotype A) protein"
FT	replace (607, C)
FT	/*tag= b
XX	
PN WO200123540-A2.	
XX	
PD 05-APR-2001.	
XX	
PF 28-SEP-2000; 2000WO-US26767.	
XX	
PR 28-SEP-1999; 99US-0156510.	
XX	
PA (GENT-) GENTEST CORP.	
XX	
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Pattem CJ;	

Sequence of human  
Retroviral vector  
Human G185V mutant  
Retroviral M4 mdr-  
Multidrug Resistant  
Human P glycoprotein  
Cynomolgous monkey  
Mutated human P-gly  
Mutated human P-gly  
Human MDR-1 DNA.  
Human ATP binding  
Human MDR-3 DNA.  
Rat multidrug resi  
Rat multidrug resi  
Rat multidrug resi  
Rat mdr1b2 (multis  
Murine multidrug r  
Human ATP binding  
Multidrug-resistan  
Multidrug-resistan  
Multidrug-resistan  
H. contortus PCP-A  
Arabidopsis thalia  
cDNA encoding mult  
A. thaliana ATPAC  
cDNA encoding the  
H. contortus PCP-A  
DNA encoding the m  
H. contortus PCP-O  
Aspergillus flavus

XX WPI: 2001-235373/24.  
DR P-PDSB: AAE00308.  
XX  
XX  
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT PGP inhibitors -  
XX  
XX  
PS Claim 9; Page 85-90; 11pp; English.  
XX  
CC The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity, by  
CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump  
CC exporting small molecules across the cell membrane. This enzyme  
CC is a member of the ABC transporter family.  
XX  
XX Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Query Match	36.6%	Score 1325.4	DB 22	Length 4279
Best Local Similarity	62.6%	Pred. NO. 0		
Matches 2190; Conservative	0	Mismatches 1256	Indels 51	Gaps 6

[illegible]

Db	898	tatacagccgaacaatttctatgtgtgvcggttcttattgtatctatgtcatatgtctct	957
Qy	759	tgctttctgtatgagaacctccctctgtattcttaatgagaaacctgtgatatcaacatcggtac	818
Db	958	ggcttctgtatagggacctccctctgtctctcccaag-----tgatatctattgtgaca	1011
Qy	819	tgctctgtctgttcttcttaagtgtaatccatagcagttatgtcatgtgaaagcaagctcc	878
Db	1012	agtaactacgtctctcttcttctgtataatattggggtcttttaattatgtgacagcatccccc	1071
Qy	879	tcactttgaaacctgcgcgaatagccggagagctgccttcaattctatcttccaggtcatga	938
Db	1072	aagcatctgaagcatttgcgaacgcaagagggagcagcttatgtaaaccttcaagataattga	1131
Qy	939	taagaaaccgcagtaataatactttctccacagctgatataaacctgataccatagaag	998
Db	1132	caataaaccgaagcattgacagctatccgaagagtgtgacataaaccagataattaaagg	1191
Qy	999	aactgtggaatttaaaaaatgttcttcttaattatccatcaagaccatctatcaagattct	1058
Db	1192	aaatttggaaattcctcaaaaatgttactactgcagttaccccttctgaaaagaagttaagattc	1251
Qy	1059	gaaaggtctgtaacctgcgaattgaagcttgagagagcaagctgcgcttgcgttcccaatgg	1118
Db	1252	aaaggtcttccaaacctgaaaggttcaagagtgtggcagacagatgtgcgtctgttgggaacagtgg	1311
Qy	1119	cagtgtggaaagatcaggtatagttccagctctctgcagaggttatatgtatccgagatgtgctt	1178
Db	1312	ctgctgggaaagcgacgcgcgctgtgcagctatgcaagaggtcttatgtgccccacagatgtgat	1371
Qy	1179	tatcatgtgtgatatgaaatgacatcacagagctttaaattgtgcgtcatctcgagaccatat	1258
Db	1372	ggtctgtatgtatgacaggaacatatagaccataaagttaagagcatctctcgtggaattatc	1431
Qy	1239	tgagagtgttgatgataagagcctgtttgttgcggaccaccatcagttaacaatatcaagta	1298
Db	1432	tgtgtgtgtgtgataaggaagcctgtgtgtgttgcaccaccagataagctgaaataactgtcta	1491
Qy	1299	tggagagatgatatgtgcctgtatgaaagagatgtagagagcagcaagaggaagcaaatgtcgta	1358
Db	1492	tggcgcgcgaaaatgtcacccatgtgatgatattggaagaagctgtttaaggaagccaatgtccta	1551
Qy	1359	tgatttatcatgtgagttctcctaataaatttaatactgttgagggaaagaaagagctca	1418
Db	1552	tgattttatcatgaaactactaactaataatttgcacactcgtgttgagagagagagggcccca	1611
Qy	1419	aatgagtgtgagggcagaagaaagagagagctgcgaattgtctgtgcctatgttcgaaccocaa	1478
Db	1612	gtctgagtgtgtgaaagaagaacaagaaatcgcgcatctgtcggggccctgtctgcgaaccocaa	1671
Qy	1479	gattctgattttatgatatgagctctacgctctgcgcctcgtgattcaagaagaacagttcacgtcttca	1538
Db	1672	gattcttctgtctgtatgtagggcaacgtcagctcttcggacactctgaaagtgaagcagtgcttca	1721
Qy	1539	agctgtcactgtgagaagacgcgagcaaaaggttcggactacaaatcgtgtgtagcaacacgcagcttc	1588
Db	1732	gtgtgcctctgtgataagccagcaaaagacgcgactaccattgtgataagcaltcgttttgc	1791
Qy	1599	tactattctgaagatgcagatttgattgtgtgacccctaaagaggtgtaatgtcgtgcggggagaaag	1658
Db	1792	tacagttctgtatgtccagatgtcatctgtgtgtttgttatgtatgagatcatattgttgagaaagag	1851
Qy	1659	agcaaatgcgtgaactatgtgcaaaacgaggtctataattatcaactgtgtatgtgcagaaga	1718
Db	1852	aaatcaatgatagaactcatgaaaggaaggaaggtcattcttccaaactgtgtcaacaatgtcagac	1911
Qy	1719	tattaaanaaagctgatata-acagatgtgagtgcaatgtacatattctatctgaaaagaaagacca	1777
Db	1912	aagaggaataatgaattgtatgagaataatgcactgtgtgatatccaaaagtgaaagatgtatgc	1971
Qy	1778	actcaactctctgtcactctgtgtgaag-----catcaagatcagactcatattgtgaaa	1829
Db	1972	cttggaaatgtctccaaaagatcacaagttcacagtttaataaanaagaaagatcaactcggag	2031



OY	1830	ggtctgaggaaatccaaccaatctaaagagabaagctt-----	1866
Db	2032	gagatatacatgccaaccaagaagccaaagacttggatatacaaaagagacttgatga	2093D
OY	1867	-----cctgaagtcctctctatataaaatttaagaagttaacaagcccgaaatggcctt	1919D
Db	2092	gaatgtaacctccaaattccctctcggagagatctcgaaagccgaactcaactcaactgagtcctta	2151D
OY	1920	tgtgtctctgaggagacttggctctctgtctcaaaatgaagactgtctatccagatattccat	1979D
Db	2152	tttgggtgttggtatattttgtgtctattataaaccggaagccctgcgaaccaagcatttcaat	2211D
OY	1980	catcttgcacaaaatatataaacatgctt--ggaataatgataaaccacataaagca	2038D
Db	2212	aatatatttcaagagataatagaatcttttaaccgagatgtagatctctgtaaacaaacagca	2273D
OY	2057	tgaatgcaagaattatattccatgatatctgtcatatttgggtgtattatttggcttggatga	2096D
Db	2272	gaatagtaaacgtttctctgtatgtttctcagtcctctggaattatctctttatatacat	2333D
OY	2097	tttaatgaagagatatttttaacgagagagaggaatttaacagatagatttaagca	2156D
Db	2332	tttccatcaagagcttcacatttggcacaagcttggggagatccctcaactaaacggtctcgata	2391D
OY	2157	cttggccttcaaaagccaatgtatataatcagaatatgtcctgtgttgaatgaagaagaaagac	2216D
Db	2392	catggttttcagatcatcagctgtgagacagagatgctcagcttggatttgatgaccttaaaaacac	2451D
OY	2217	cacagagagctttgacaaacaatatltagccatagatatagcacaaattcaagaagcaaacg	2276D
Db	2452	cactggagacatttgcacaacacagcgtctgcacatgattgcgtctcaagttaagaagggctatag	2511D
OY	2277	ttccagagatttgcgtcttcaacacaaatgaaataacatgaggaacttccagatcatcttc	2336D
Db	2512	ttccagagcttgcgtctcaattcccaagaatatagcaaaactcttggacagcgcatattatatac	2573D
OY	2337	ctttatatatgabatvgggagatgtacatctcgtatcttcgaglatgtctccagatcttgcgct	2396D
Db	2572	cttaatcatatgttggaaatttaaacactttactctttagaatttgaatccatcatcatgcaat	2631D
OY	2337	gacaggaatgtattgaaacccgcagcgaatgactgtattgcccaacaaagataagcaagaact	2456D
Db	2632	agcgaagagttgttgaataatgaataatgttgtctgtagacagcactgaaagataagaagaagact	2693D
OY	2457	taagcatgtctggaagaagatagcaactgaagcgtttggagaatatagctatagatgtgta	2516D
Db	2692	aggaagagcttgggagaagcttgcctacagaagccacgcaaaaactccgaactgttgttctct	2751D
OY	2517	aacaaaggagaaaaacccctctcagacaaatgtataatgaaagatcgtctcagactcaacaaga	2576D
Db	2752	gactcgggagagagaagtttgaatacatgtatgacaacgaagatttgcagaatgacataaga	2811D
OY	2577	tacctcgaagaagaacacagatatatttggaaagctgtatgtacatccagcatgcttatata	2636D
Db	2812	ctctcttgagagaagcacacacatcttcgggtgtcattcttctatccaccaggaatgatgta	2873D
OY	2657	tttggctatagcagcagaggttctgatttggagcctatctaatacgaactgcgaagatgac	2696D
Db	2872	tttttctctatgtcgtgcgttccggtttgtgtgtcctactcttggcgaagaaagatcatbaa	2933D
OY	2697	cccaagagagcgtgtcatagtttctaactgaatcttgatataatgagcatatgagccatgaa	2756D
Db	2932	cttccagagatgcttcttggatcttccagctatgtccttggctgcacatgacgttgggca	2991D
OY	2757	aacgctcgttttggctctcgtgaatatcccaagccaaatcgggggctgtgcgcatgttgc	2816D
Db	2992	ggtcagttcatcttctcgtactatgccaagccaaagatatacgaagcccaagctatcat	3051D
OY	2817	cttgttggaaaagaacccaatatagacagccgcgcatgaagaaggaagaaagcagacac	2876D
Db	3052	gatatcttgaanaaagccctctgtatgtacagctacagccctcagccgtctaaagccaataac	3113D

QY	2877	atgtaagagggaattagaaatttcgaagaagctctctctctctatccatgtgcgcagatgct	2936
Db	3112	gttcgaagaagaaatgtgacatttaataagtgctgtgtccaactaccatccacctgcacagacat	3171
QY	2937	tttcaatcccggtggtctatccctcagttatctgagcgaggaaaggaacgaatgagcttgggg	2996
Db	3172	ccccgtgtcccaagggtctgagccctcgagggtgaagaaggccgaacgcgtgcccctctag	3211
QY	2997	gagcagcgccgtcttggaagaagcattctgtccaactctgcagacgaactttagaccgt	3056
Db	3232	tagcaggtgctgtggtggaaggaacagatgtgtcaagctccctcagaagcgctcttatgaccctt	3291
QY	3057	gcaaggaacagatgctgtctgtatggtgtgtgaatgcacaagaattgaatgtacagtggtccg	3116
Db	3292	ggctgtgtctcagtgctaatgtatgtggcaagaagataaagcacctgtaatgtccagtggtccg	3351
QY	3117	ttcccaatagcaatggttctcttaagaagccttggtgccttccaactcagatgtgcgtga	3176
Db	3352	agcaaaccttggtgcctgtgtctcaagggtcccaactcctgtttgactgtgaagatgtccga	3411
QY	3177	catcgccatgtgtgaacaacgcgtgtgtgtgtgcattagatgtgaatcaagaagcgccga	3236
Db	3412	catgtgctatgtgaacaacgcgtgtgtgtgtatcaactgtgaagatgtatgtcagcgccaa	3471
QY	3237	tgcagcaaatatccattctttatgtgaaggtctccctgtgaataatacacacaagaatgtg	3296
Db	3472	ggaagcccaacatatcacacttcatctgcagacactcctcgtgaataatacaacacagtag	3531
QY	3297	actgaagaaggaacagcttctctgtgcgcgaagaacaagaactagctatgtcaaggtcct	3356
Db	3532	agacaagaagaaaccagctctctgtgtgcgaagaacagcgatgtccatagctgcgcctc	3591
QY	3357	tctccaaaacccaaattttatgttgaatgtgaatggccacttcagccctcgataatgtac	3416
Db	3592	tgtttagaagactatatttctgttcttggatgtgaagctacatcagctctgtgaatacgaag	3651
QY	3417	tgaagaagtggtgtcagcatgcccctgtgaataagccagagcggaagacatgtcctagtg	3476
Db	3652	tgaagaagttgttccaagaagccctgtgcagaagcagagaagcgccgacctgtatgtat	3711
QY	3477	cactcacaggtctctctgcaattcagaacgagaatttgatagtggtctctgcaaatgga	3536
Db	3712	cgccacacgcgtctgtccacatccagaatgacagatttaatagtggtgttcagaatgca	3771
QY	3537	gataaaggaagaaggaactatcaacaagctcccgagaatactgcagacatatatttaagt	3596
Db	3772	agtcagaagagcatgtgcacaatacaacagctgcgtgcgcgaagaagcattatatttccat	3831
QY	3597	agtgaaatgcacagtcag	3613
Db	3832	ggtcagtggtccaggtg	3848
RESULT 2			
AAD03505			
ID	AAD03505	standard; cDNA; 4279 BP.	
AC	AAD03505;		
XX			
DT	13-JUN-2001 (first entry)		
XX			
DE	Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.		
XX			
KM	Dog; P-glycoprotein allelic variant; PGP: multidrug transporter;		
KM	MDR1; drug bioavailability; transgenic animal; genetic model; ss.		
XX			
OS	Canis familiaris.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	17..3862	
FT	/*tag= a		
FT	/product= "Dog P-glycoprotein (PGP) allelic variant		
FT	(Genotype B) protein"		

FT	allele	replace (91, T)
FT	allele	/*tag= b
FT	allele	replace (507, C)
FT	allele	/*tag= c
XX		
PN	MO200123540-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000OWO-US26767.	
XX		
PR	28-SEP-1999; 99US-0156510.	
XX		
PA	(GENT-) GENTEST CORP.	
XX		
XX	Stocker PJ, Steime1-crespi DT, Crespi CL, Relf TC, Patten CJ;	
DR	WPI: 2001-235373/24.	
XX	P-PSDB; AA00309.	
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful	
PT	for determining the bioavailability of drugs and for screening for dog	
PT	PGP inhibitors -	
PS	Claim 9; Page 93-99; 11pp; English.	
XX		
CC	The invention relates to dog P-glycoprotein (PGP) also referred	
CC	as multidrug transporter (MDR1) and nucleic acids encoding them.	
CC	The invention also includes fragments and biologically functional	
CC	variants of dog P-glycoprotein. PGP and their nucleic acids are	
CC	useful for determining the bioavailability of drugs and for	
CC	screening PGP inhibitors. They are useful for the diagnosis and	
CC	treatment of conditions characterised by PGP activity, by	
CC	reducing or increasing PGP activity in a cell. PGP nucleic acids	
CC	are used as oligonucleotide probes. Complements of PGP nucleic	
CC	acids are useful as antisense oligonucleotides, to induce a PGP	
CC	'knockout' phenotype. They are used to prepare a non-human	
CC	transgenic animal, which are valuable as genetic models for	
CC	human diseases.	
CC	The present sequence is dog P-glycoprotein (PGP) allelic variant	
CC	(Genotype B) cDNA. The PGP enzyme functions as an efflux pump	
CC	exporting small molecules across the cell membrane. This enzyme	
CC	is a member of the ABC transporter family.	
XX		
SQ	Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;	
	Query Match	36.6%; Score 1325.4; DB 22; Length 4279;
	Best Local Similarity	62.6%; Pred. No. 0;
	Matches 2190; Conservative	0; Mismatches 1256; Indels 51; Gaps
QY	162 caggttaccctgtatattgtgaatggtgtgctgctgatttgyttacataca	221
DB	11 11 11 1111 11 1111 11 11 11 11 1111 11	
DB	358 cagatagctcatattatacagtggtgctgtgctgctgctgctgtacatcca	417
QY	222 gatttccttgygattataaactgcagcacgacagacaaaggattcgaaaacgtttt	281
DB	11 1111 11 11 11 11 11 11 1111 11 11 1111 1111	
QY	418 ggttcatcttgcgtgcctgcagcaggaagacatcctcaaatatagaacaatttt	477
DB	11 1111 11 11 1111 11 1111 11 11 11 11 1111	
QY	282 tcatltaagtttggcagcagaacatgcgcgtgtttgtaactctgtgcatggygaactaa	341
DB	11 11 11 11 1111 11 1111 11 1111 11 11 11 1111	
QY	478 tcatgtaatacagcagcagagatggtgcgtgttgcgtgcatgaagttggygaactaa	537
DB	11 11 11 11 1111 11 1111 11 1111 11 11 11 1111	
QY	342 cactcgatcagcagaacatcgtt--acaaaatcagttggtlaltggaataaagattgctc	398
DB	11 11 11 1111 11 11 1111 11 11 1111 11 11 11 1111	
QY	538 caccgcgtccacagcagaatgctcccaaaatcaatgaagaatgtgcgaacaaattggaat	597
DB	11 11 1111 11 11 11 11 11 11 11 11 11 11 11	
QY	399 gttgtttcaaaaacatgctactcttlttgcattggtcgtgcagattggttgygaaggtcg	458
DB	11 11 1111 11 11 11 11 11 11 11 11 11 11 11	
QY	598 gttcttcaatcaatagcaacatttttaacgcgttttatagtggttttaacagtggtg	657
DB	11 11 1111 11 11 11 11 11 11 11 11 11 11 11	
QY	459 gaactcaccctatgtaactctatccagctctcctcttataatggtcttgagcgacgcatg	518
DB	11 11 1111 11 11 11 11 11 11 11 11 11 11 11	

Db	658	gaacgtcaacctgtgatttlttgccatcagccctgttcttgaaactttagccgcacatctg	717
Qy	519	ttccaggatgtaacatccattgacccgaatgaagaaattgacatcttccaaagcttgggc	578
Db	718	ggcacaagatactatcttcattctatgataaagaacctgtgacctatgtgaagaactggagc	777
Qy	579	tgtgacgaagaagcttctgtcatcatccgaagatcatagcctttaaggcccgagaa	638
Db	778	agtagctgaagaagctttagcagcaatcagaactgtagtcccttcttgagagcaaaagaa	837
Qy	639	agaactccaagatgatacacagaatctcaaaagtgaagaagatttggcatnaaagagac	698
Db	838	agaacttgaagaagtacacaanaaatltagaagaagcttaaggaatttggataaagaagaac	897
Qy	699	tatatgtccaagaatgtgctctgtgtgctgtgtacatcctttagaattgaaactatgact	758
Db	898	tatacggcccaacttcttatgttggcgttcttattgactatgatacatatgacct	957
Qy	759	tgtctttgtatgaaacctccttgaattcttaatgagaacctgtatatccatccggagac	818
Db	958	ggcttctgtatggagccctcctgtgcctccacg-----tgaatatctatgtgaca	1011
Qy	819	tgtctgtcgttcttctttagtgaatccatagcagtatgttcatcttggagcgacgtctc	878
Db	1012	agtaatcaactgtcttcttcttctgtatnaattgggctttagtatatggacagaatcccc	1071
Qy	879	tcaactttaaacctctgcgaataagcccgagagcgctgccttcatatttccaggttatga	938
Db	1072	aagcatitgaagcatttgcgaacgcgaagagagcgagcttatgaaatctcaagaatatga	1131
Qy	939	taagaacaaccagtatagataactttccacagctgtgatatataaactgtaatccatagaag	998
Db	1132	caataacaacaagcatttgcagcattatcgaagagtgtgacataaacaagataaattaaagg	1191
Qy	999	aacgttgaattttaaanaatttcttcttccaatttccatcaaacacatcatcaaatctc	1058
Db	1192	aaattctgaattccaanaatttgcatacttcagttaaccccttcgaaaagaagttaagatctc	1251
Qy	1059	gaaaggtctgaatctcgaatltaaagctctggagagaacagctgcgttgtcgttccaatg	1118
Db	1252	aaaggtctccaactggaaggtttcagagtgtggcagacagctgtgcgttgttggtaacaatg	1311
Qy	1119	cagttggagaagatcacgttagtccaagctcttcgcgaagaagtatatatgacccgagatg	1178
Db	1312	ctgcggggaagacacgcagctgtgcagctgtgcagagagctctatagaccaccaagatgcat	1371
Qy	1179	tatcatgtgtgatagaatgtagacatcaagaagattttaaagttgtgcgcattatcgaagacat	1238
Db	1372	ggtctgctatttgaatgacagagacatltagaccataaagttaaggcatcttcgcggaattac	1431
Qy	1239	tggagttgttagtccaagaagcctgttctgttcggagaccaacatcagtaacatacaagta	1298
Db	1432	tgtgtgtgtgtagtcagagcgctgtgtgttgcacacacagatagctgtgaanaaatctgtca	1491
Qy	1299	tggacgagaatgattgtactgaatgaaagagatggagagacgacaagggaagcaaatgtgcta	1358
Db	1492	tggccgcgaaaatgtccacatctgtagatcttgagaagacgctttaaggaagccaatgtgcta	1551
Qy	1359	tgaatttcatctgagtttcccttaataatttaatacatltgttagtggggaanaagtgactca	1418
Db	1552	tgaatttatcatgaactacactaataatttgcaacctcgtgttgagagagagggggccca	1611
Qy	1419	aatgagttgagagtgacgaanaacagagatcgaattgctgtgtccttgaattcgaaccccaa	1478
Db	1612	gcttgagttgtgacacgaanaacagagaatccgaatctgcgtggccctgtgttcgcgaacccaa	1671
Qy	1479	gattcttgatttagatgggttaagcttgcctgcctgcgatttccgaagaagaagtccaagctgttca	1538
Db	1672	gattctctgcctgtgtgtgggcaaacgtccaagctctgtgaacactgaaagctgtgttca	1731
Qy	1539	agcttcaactggagaagtgaggaacaaagttcgagatacaatgtgtgtgacacacgcgaatctc	1596
Db	1732	ggttgacctgtgataagccacgaanaaagccggaactacacattgtgatagtctcatcgtttgc	1791

1599 tactatcgaagtgacagatttgatctgacccctaaaggatgaatgcgtcgagaaag 1658  
1792 tacagttcgttaatgcgcgtatgcatctgctggttttgatgtagtgcatctgtagaagaa 1851  
1659 agcacatgctgaactaatgagcaaaacgagtcataattactctgtgatgtccagaa 1718  
1852 aatcatgatgaactcagaagagaaggacattactccaactctgtccacatgacagac 1911  
1719 cattaaaagctgtatga-acagatgagtgcaatgacatattctactctgaagaaagacca 1777  
1912 aagagaagaatgaatgtgtagaagacagctggtgaatccaaagtgaagtgatgc 1971  
1778 actcactccctgctgacctgtgaaagag-----catcaagtcagacttattgacaa 1829  
1972 ctggaagaatgctcccaaaagatccagggtccagtttaataaagaagatcaactcgacg 2031  
1830 ggcgtgagaatcccaactcaaaagagataagtcct----- 1866  
2032 gagtatcatgaccccaagagcaagagacgaagagctgtgatacaaaagagacttgatga 2091  
1867 -----ccgaagctctctatataaaatttaagttaaacagcctgaatgccttt 1919  
2092 gaatgtaacctccagttctctctctgtagagattctgaagctgaactcaactgaatgcctta 2151  
1920 tctgtctctgaggaactggtctctctcttaagtgaactgttcaatcagatatttccat 1979  
2152 ttctgtgtgtgatatatttgtgtcattataaagcgaagcctgacacagcctttcaat 2211  
1980 catcttgcacaaatataaaccatgltt--ggaataatgataaacaacacatlaaagca 2036  
2212 aatatcttaagagattatagagatcttaaccgagatgagagctcctgaaacaaacgaca 2271  
2037 tgatgcagaattttctccatgatatatcgtcatcttggtgttatttgccttgcaatga 2096  
2272 gaatgtaacagatttctgtatgttctcagttcccttgaaattacttcttatttaccat 2331  
2097 ttctcagagagattattttacgagagcgaagggaatttttaagatgagattgaaga 2156  
2332 ttctccagaggtccacatcttgcaaaagcvggagagaccctcaaaacggtcttcgata 2391  
2157 ctgagcttcaaaagcatgltatatacagatatctgcgtgttgaatgaagaaagaaacag 2216  
2392 catggtttcagatccatgctgtagacagatgtagcgtgttgaatgagcccaaaacac 2451  
2217 cacagagagcttgacaacaatatagccaatagacatatagcaaaatccaagagcaacag 2276  
2452 cactggaacatgacaacacagcgtgccaatgtagcgtctcaagttaaaagggctatag 2511  
2277 ttccaggttgagctctaaacaaatgcaactacatcagatggacttcaagtatacttcc 2336  
2512 ttccaggttgcgtcgtcatataccagaataatagcaaatcttggaacagagcatattatac 2571  
2337 ctctatatactgagtgagatgacattccgtatctctgtagtattgctccagtaactgccc 2396  
2572 cttaactcatggttggaacttaaacacttttactcttagcaattgtgataccatcttgcaat 2631  
2337 gacaggaatgattgaacccgacgaatgactgagatttgcacaaagaagaacgaagact 2456  
2632 agcagagagctgtgaaatgaaatgtgtctgacaagacatgaaagaataaagaagct 2691  
2457 taagaatgctggaagatgacacgtgaagccttgtagagataatagctactatgttcaat 2516  
2652 aagaagagctggaagatgctlaacagagccatcgaaaactccgaactgtgttctctt 2751  
2517 aacaagggaaaaagccttcgagcaaatgtatgaagagatgcttcagactcaacagaa 2576  
2752 gactcggagagaaagatttgatacatgtatgacagagatttgcaagctacccatacagaa 2811  
2577 taactcgaagaagaagcagatattggaagctgtatgcaatcagcactgacctatata 2636  
2812 ctcttggagaaagacacatcttcgggtctcattctatccaccacagcaatgattga 2871

2637 ttctgcctatgcagcagggtttcgaatttgagacctatattaaatcaacgtgacgaatgac 2696  
2872 ttcttcctatgctgctgtttccggttttggtgcttactcttggtggaatgattgataa 2931  
2697 cccagaaggcatgctcatagtttctactgcaattgcatatgagctatgagccatcgaaa 2756  
2932 ctctcagaatgcttcttgatattctcagctatgtcttctgtgcatgagcagtgaggca 2991  
2757 aagctcttctggtctcgtatattccaaagccaaatccggggtcgtgcactgtttgc 2816  
2992 ggtcagttacttctgctgactatgccaagccaaagtatcagacagcccaagtcataat 3051  
2817 ctgtgtggaagaaagcaacaaatatagacagccgagtcagaagagggaagaaagcgaac 2876  
3052 gatcatgaaaaaagccctctgattgacagctacagcctccacggtccctcaagccaatac 3111  
2877 atgtgaagggaatttagagtttcgaagaagctcttctctataatcctatgctgcagatgt 2936  
3112 gttggaaggaatgtagcaatttaagtgctgttcaactatccactcgacagacat 3171  
2937 ttctactctcgtggtctatccctcagatattgagcgaaggaagacagatgtgtgg 2996  
3172 ccccggtctccaggggtctgaagctcgaagtgagaagggccagacgctggtccctgtag 3231  
2997 gagcagcgtctgtaggaagaaagcactctgctcaactctgcagagacttaatgacccgt 3056  
3232 tagcagtggtctgtggaagagacagatgttgcagctctcagagcgtcttctatgacccct 3291  
3057 gcaaggaacagtgctgtttagatgtgtgtagtgcacaaagaatgaaatgatacagtgctcg 3116  
3292 ggtctgtcagtgcaatattgattgagcaagaagataaagcaccgtgaatgtccagtgctcg 3351  
3117 ttcccaatagaaatcgttccctcaagagcgtgtgtcttcaactctgcagcatgtcgagaa 3176  
3352 agcacacctgggcatcgtgtctcagaagcccaactcctgttgactgcaagattgcgcgaa 3411  
3177 catgcctatgtagacaagacgcgtgtgtgcatatagatgagatcaaaagacgcgcgaa 3236  
3412 catgtcctatgtagacaagacgcgtgtgtatcacaatgaaagatatagcaagcccaa 3471  
3237 tgcagcaaatatcacttcttatttgaaggtctccctggaataatacaacacaaagttgg 3296  
3472 ggaagccaacatacaacacttcatcgagacactcccttggaataatacaacacaaagttag 3531  
3297 actgaagagtagacagcttctcgcgcgcagaaacaaagacagctatgtcaagggctct 3356  
3532 agacaaaggaacccagctctcgtgtgcagaaacagcgcatttgcatagctgcgcctct 3591  
3357 tctccaaaacccaaatttatttatttgaatgagccacttgaagccctcctgataatgag 3416  
3592 tgttagacagcctcatatttgccttcttgaaagctacatcagctcgtgatacagaag 3651  
3417 tgaagaagtgtgtcagcagctcttgataaagccaggaaggggaagagacatgctatgtgt 3476  
3652 tgaagaagtgtgtcagaagagcccttgacaagccagagaagggccgacactgattgtgat 3711  
3477 caatcaaggtctctctgcaatcagaacgcaagatttgatagtggtctgcacaatgga 3536  
3712 cgcacacgctgtgtccacatccaaatgcaatgcaatttaatgagtggttccagatgga 3771  
3537 gataaaggaacagaagacatcagaagctccttggaatctgaagcatabatttaagt 3596  
3772 agtcaagagacatgacacacacacagcgtcgtgcccgaagaagcatabatttccat 3831  
3597 agtgaatgcacagtcag 3613  
3832 ggtcaggttccagctg 3848

RESULT 3  
AAD03506  
ID AAD03506 standard: cdna: 4279 BP.  
XX

[illegible]

Db	358	cagctatgcattatattacagctggagctggctggcgtggcgtggcgtggcgtggcgtttacatcca	417
Qy	222	gattccctcttgagatttaactctgacgcagcagacagacaaagagatctcgaacagctttc	261
Db	418	ggttcattctcgtggccttgccagcaggaagacagattactcaaatltagaanaacatttc	477
Qy	282	tcaatcagttttgacagagacatcgctggcttgatagcgttgacatcggltgaactaa	341
Db	478	tcatgctatctctgacaagagagatctgcgtgttttgacgtgacgtctggaggacttaa	537
Qy	342	caatcgatcgacacactt---acaaatcagttgattgttatggaataagatctgctc	388
Db	538	caccgcgctcacagacgattgtccccaataatgaagaatctggcgcgaatcttggat	597
Qy	399	gttgtttcaaaaacatgctactcttttcgatctggcctgbcagltgtgttggtgaaggctg	458
Db	598	gttcttcaatcatagacaacttttccacgggttttattagtggggttttacaagctggltg	657
Qy	459	gaactcaacctagtgactctatccacgctcctcttataagtgcttaaggaggagacgt	518
Db	658	gaagctaaccccttgatttgattctggcgcacacgcccctgttcttgactcttcagccgcac	717
Qy	519	ttcttagagctgtcatccatcttgacccagttgaagattaaagtgtcctcttccaaagctgg	578
Db	718	ggcaagaactactactctcaatttacttgataaagaaccttggtcctatgtcaaaagctgg	777
Qy	579	ctgtgcacaaagaagctctgtgcatcaatccgaaacagctacagctcttbaaggccagga	638
Db	778	agtagctgaagaagcttcttagcagaatcagaactggtatgctcttgaggagcaaaagaa	837
Qy	639	agaacttcaaaagtatacacagaaatccaaagctgcaaaagatttttgcatataaagaag	698
Db	838	agaacttgaagaagtgatacaaaaatttagaagaagctaaaggaattggattaaagaagc	897
Qy	699	tatagcttcaaaagctgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	758
Db	898	tatcaaggccacaattcttatgtgtgcgcttctcttatgtatctatgcatctaatatgctc	957
Qy	759	tgcttttttgatgtagaacctccttgatcttctaattgaagaaacctgatatccatcgga	818
Db	958	ggcttctcgtatggagacctcctgtgctcctccag-----tgaatatcatattgaca	1011
Qy	819	tgctctgctgtttcttctttagtgaatccatagcagttatgtatgggggcagcgctcc	878
Db	1012	agttactcactgtcctcttcttcttgatataattgggcttttagatattggcagcgctcc	1071
Qy	879	tcaatttgaacacctgcgaatagccggaagagctgaccttcatatttccaaagtattga	938
Db	1072	aagcatttgaagcatttgcgaacgcgaagaggaagcagcttagaacttccaagaatttga	1131
Qy	939	taagaacccgatatagataacttccacagcgtgatataaacctgaatccatatagaag	998
Db	1132	caataaaccaagcatgtgcagctatctcgaaagtgtagacaataaccgaataatataag	1191
Qy	999	aactgtggaatttaaatgtttcttccaattatccaataagccatccataagaatctc	1058
Db	1192	aaatttggatccaanaaaatgttcaacttcagttacaccttcgcgaagaagttaagaactc	1251
Qy	1059	gaagaagctgcaatccagaaatcttaagctctggagagacagctgccttgctgcgtcca	1118
Db	1252	aaaggtctccaaccggaaaggtctcagatgtggcagacagctggcgtgtgtgggaaga	1311
Qy	1119	cagctgggaagatgaacgtgagctccagctctctgcagaggttatatgatccggatgat	1178
Db	1312	ctggcggaagagcagcagcgtgagcagctgagtcagagagctctatgacccaagatgcat	1371
Qy	1179	tatcatgctgagatgagaatgacatccagcgttataatgtgcgcatattacaggaacata	1238
Db	1372	ggtctgtattgttgacagcagcatatagacataatgataagcatctcttcggaaattac	1431
Qy	1239	tggagaggttaataaaggccgtgttttttggaagacacatcaagtataaatataagta	1296
Db	1432	tgggttggttggaatccaggaacctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1491

OY	1299	tgcgcgagatgacgtgacgtacgtacgaagatgtagaagacgacgaagcgaacatgctga	1358
Db	1492	tgcgcgagaaaatgctccaccctgtagatgtagaagaagctgttaagagaaagccatgctcta	1551
OY	1359	tgaattatcaatgtaggttccctaataattatatacatgtgtaggggaaaaagagctca	1418
Db	1552	tgaattatcatgtaaacctactaatataattgtgacactcgtgtgtgagagaggggccca	1611
OY	1419	aatgagtgagagggagaaaagagaggtctgcaattgctctggtccttagtttgtaaacccca	1478
Db	1612	gcctgagtcggtgagaaaagaaaacagagaatctgcacattgctcgggccctggttgcgaacccca	1671
OY	1479	gattctgattttagatgtagagctacgtctcgtccctgtagatltcaagaagcgaatcagctgtca	1538
Db	1672	gattcctctcgttgtagtagaggccaacgtcagctcgtcgtgacactgtgaagtgaaagctgttca	1731
OY	1539	agctgcacatgtagaagcgagcgacaaagctcggagctacacatcgtgtgtagcaaccagacttc	1588
Db	1732	ggctgcctctgtagatagagccagaaaagcgcgactacacattgtagatgactcatcgttttgc	1791
OY	1599	tacattctcgaagtgacagatttgaatttgacccccaaggtgaaagtcgtgcggagaaag	1688
Db	1792	tacagcttcgtagaagcgaatcgtcatcgtcgtgtttttagatgtagagtcattgtgtgaaaag	1851
OY	1659	agccatgctgcgaactaaatgtagcaaaaacgaggtccatatatttcaactgtgattgtagcaaga	1718
Db	1852	aatcatgatggaactcatgtagaagaagagcgatttcaacttcaacattgttcaacatgtcgaa	1911
OY	1719	tattaaaaaagctgtaga-acagatgtagatcaatgacatatctctcgtgaagaagaccca	1777
Db	1912	aagaggaatgtagaaattgtagtatagaaaatgccaactggtgatccaaagatgtagaagtgc	1971
OY	1778	actaactccctcggcgaactcgtgtagaag-----catcaagtcgaagctcatgtgaa	1829
Db	1972	cttgagaaatgctctccaaaagatccaggtgcgcagttataaaaaagagatcaaccgtgag	2031
OY	1830	ggctgcggaatcccccacaactcaagatagaataagctt-----	1866
Db	2032	gagtatcatatgcaccacaaagcgcaagacagaagaagctgtgtacaaaagagactgtga	2091
OY	1867	-----cctgaagctctctctatataaaaaatttaaagttaaacgaagcctggaatgctctt	1919
Db	2092	gaatgtacctccagatttccctctcgtgagagatctcgaagctgaaactcaactgaatgctcta	2151
OY	1920	tgtgtcttcgtggagcatctgctctcgtttcctcaaatggaactcgtctacagatatttccat	1979
Db	2152	tttctgtgtgtgtatatatttctgtctattataaacggaagcctgcgaaccagcattttccat	2211
OY	1980	catctctgcgaaaaatatataaccatgctt--ggaataatgataaaccacataaagca	2036
Db	2212	aattatttcaagagttatagggatcttcaaccgagatgtagatctctgtaaaaacagcca	2271
OY	2037	tgaatgcgaatatattatccatgtagatattgcatattgggtattgtatttgccttgta	2096
Db	2272	gaatcagtaaacgttcttcgtatattgtcttcagcccttggaaattatctcttattaca	2331
OY	2097	tttcatgcagagattattttagcgcagagcgagcggaagattttaaagatgagattaagca	2166
Db	2332	tttctctcagaggtctcacttcttgtagcaagcgtggggagatccctcactaagcgtctcgta	2391
OY	2157	cttggtccttcaaaacatgtagtatatacaggaatctgcgtgtttagatgaaaagaaacag	2216
Db	2392	catgttttctcagataccatgtctgtagacaagatgctcagctgtgttagtagaccttaaacac	2451
OY	2217	caacggagagctctgtagcaacaatattagcgaatagtatagcaacaattcaagggcaacgg	2276
Db	2452	caatgcagacatgtagcaaccgagctgcgcacatgtagtcggtccagattaaagggtctatag	2511
OY	2277	tttccagattggtcgtcttcaacacaaatgtaaacatcgtggagcttccagttcatcatctc	2336
Db	2512	tttccaggtctgtcgttaattacccagatatagaaaatctctgggaacgaggtattatata	2571

QY	2337	cttataatagatvggagatgscatccctgatalctgtagtactgccaagtactgacct	2336
Db	2572	cttaatactatgcttggccaattaaacctttactctttagcaattgtacccaatcatgcat	2631
QY	2397	gacaggaatgtattgaaccccgacgaatgcatgctgatttgcacaagaagttaagcaagact	2456
Db	2632	agcaggaatttgttaaatgtgtaaatgttgccttgcgaagaagactcgtaagaataagaagagct	2691
QY	2457	taagatcatcttggaagaataagtaacaactgaaagcttttggagaatacagtaactatgttcaat	2516
Db	2692	agaagagagcttgggagaagatttgcatacagaagccatcgaaaacttcgcagactgtgtttctt	2751
QY	2517	aacaagggaanaagccttcgacgaataatgtatagtagaagatgcttcagactcaacaganaa	2576
Db	2752	gacttcggagacgaagaattttgaatacatgtatgacagagatttgcagaattacccaatacaga	2811
QY	2577	taacctggaagaacacagaatttttgaagctgttataatccagcagatccattata	2636
Db	2812	ctctcttggagaagacacacacttcctgggtctcatcttcatcacaccaggaatgaatga	2871
QY	2637	tttttcacatbcacaagggcttcgatttgcatttgcagacctataatccaagctggaagatcac	2696
Db	2872	tttttcacatbcctgctgcttccggttttggctactctgggcgaataatgattcatga	2931
QY	2697	cccaagaggcatgttcatagtttttaactgcaattgcataatggagctatgcccacgaaa	2756
Db	2932	ctttcagaatgttcttcttggatattctcaagctatgtcttcttggccaatgacgtgggca	2991
QY	2757	aagctctgttttgcctccctgaatatcccaagccaaatccggggctgctgcacatctgttc	2816
Db	2992	ggtcagttcatcttctcctcactatgccaagcgcaagaatatacgaagccccaagttcatat	3051
QY	2817	cttgttggaaaagaacacaaatataagacacgcatcaagaagggaanaagcaaacagac	2876
Db	3052	gatactgaaaaaagccctctgatttgaacgctcacagccctcaagcctccaagccaataac	3111
QY	2877	atgtgaagaaggaaatttaagatttctcgagaagctcttctctctatccatgtcgcccaagt	2936
Db	3112	gttggagaagaatgtgacatttaatgaaggtcggttccaacttccacatccacctcgacagacat	3171
QY	2937	tttaactctccgtgttataccctcagtatgtgacgagaagaagaacagtatgcatgttgg	2996
Db	3172	ccccgtgctcccaagggtctgcagccttcgagtgagaagaaggccagacgctggccctctag	3231
QY	2997	gagcagcgcgcttggggaaaagcaactctgttcaactctgcagagagacttaagaccctg	3056
Db	3232	tagcagtgvgctgvtgggaagcaagaatgttcttccagcctccagaagcgctcttaagaccct	3291
QY	3057	gcaaggaagaatgtctgttcttaagtgvtgtagtgcanaaagaattgaaatgtacagttgctcg	3116
Db	3292	ggctggttcagtgctaatgtatgtgcgaagaagatataagcaacctgaatgtccagttgctcg	3351
QY	3117	ttcccaatagcaatcgttccctcaagaagctgtgtctctcaactgacgacatgtctgagaa	3176
Db	3352	agcaaaccttgggcatctgtctcagggagccacatccgttttgaacttgcagatgtgcgaga	3411
QY	3177	catgcgcttatgtgaacaacgctgtgtgtgtgcataatgataagatccaagaagcgcgcaa	3236
Db	3412	catgtccctatgtgaacaacagcgctgggtgcgtatccatcagatgaagatgtgtgcagcgcaa	3471
QY	3237	tgcaagcaaatatccatcttcttatttgaaggtctcccttggaaatacaacaacaaagttcg	3296
Db	3472	ggagggcacaactataccaacttcatcggagaacatcccttggaaatacaacaacaaagttcg	3531
QY	3297	actgaagaagacagcttctctggtcggcgaagaacaagaactagctatgtcaagggctct	3356
Db	3533	agacaaagaacccaagctctctctgttggccagaacaacagcattgtccatagctctcgctct	3591
QY	3357	tctccaanaaacccaattttatgttgtgtgtggtgcacatttaagccctctgataatgaac	3416
Db	3592	tgtttagaacgctccatatttctgtcttttggatggaagctacatcaagctctgtgatacagaag	3651
QY	3417	tgagaagaattgttcaagatccctctgtataagccaggaacgggaaggaatcgtcctagtgt	3476

Db	3652	tgaaagaggttgcacaaagaagccctgcgaacaagccagaagaagccgacactgcattgcat	3711
Oy	3477	cactcacaggctcctctcaatttcagaacgcaagattgtagtgcgtcgcacaatgga	3536
Db	3712	cgccacacgctgtccacacacatccagaatgcagatttaagtggtgttcgaatggaa	3771
Oy	357	gataaaggacaacgaacatcatcaagaagctcctgagaacatcgacacatatattga	356
Db	3772	agtcagaagagatgacacacatcaacagctgtcgtccagagaagcatcatatttccat	3831
Oy	3597	agtgaatgcacagtcag	3613
Db	3832	ggtcagttccagctg	3848
RESULT 4			
ID	AAD03488	standard; cDNA; 4279 BP.	
XX	AAD03488;		
XX	13-JUN-2001	(first entry)	
XX	Dog P-glycoprotein (PGP)	cDNA #1.	
XX	Dog: P-glycoprotein; PGP, multidrug transporter; MDRI;		
KW	drug bioavailability; transgenic animal; genetic model; ss.		
XX	Canis familiaris.		
OS			
XX	Key	Location/Qualifiers	
EH	CDS	17..3862	
FT		/*tag-a	
FT		/product="Dog P-glycoprotein (PGP) #1"	
XX	WO200123540-A2.		
PN			
XX			
PD	05-APR-2001.		
XX			
PF	28-SEP-2000; 2000MO-US26767.		
XX			
XX	28-SEP-1999; 99US-0156510.		
PR			
XX	(GENT-) GENTEST CORP.		
PA	Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;		
PI	WPL: 2001-235373/24.		
DR	P-PSDB; AAE00303.		
XX			
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful		
PT	for determining the bioavailability of drugs and for screening for dog		
PP	PGP inhibitors -		
XX			
PS	Claim 3; Page 58-65; 11pp; English.		
XX			
CC	The invention relates to dog P-glycoprotein (PGP) also referred		
CC	as multidrug transporter (MDRI) and nucleic acids encoding them.		
CC	The invention also includes fragments and biologically functional		
CC	variants of dog P-glycoprotein. PGP and their nucleic acids are		
CC	useful for determining the bioavailability of drugs and for		
CC	screening PGP inhibitors. They are useful for the diagnosis and		
CC	treatment of conditions characterised by PGP activity, by		
CC	reducing or increasing PGP activity in a cell. PGP nucleic acids		
CC	are used as oligonucleotide probes. Complements of PGP nucleic		
CC	acids are useful as antisense oligonucleotides, to induce a PGP		
CC	'knockout' phenotype. They are used to prepare a non-human		
CC	transgenic animal, which are valuable as genetic models for		
CC	human diseases.		
CC	The present sequence is dog P-glycoprotein (PGP) cDNA. This		
CC	sequence is also referred as Genotype C cDNA. The		
CC	PGP enzyme functions as an efflux pump exporting small molecules		

[illegible]

QY	1119	cagcggagaaagatcagcgttagtcacgacttcttcgacagaggttatatgatctccgatatgctt	1178
Db	1312	ctcgccggaaagagacagaccgttcgacgtctgtagagaggtctctatctggccccacagatgagcat	1371
QY	1179	tatcatcgtctgagatgagatctacacgcagctttaaattgctggcgactctctcgagaccat	1238
Db	1372	gctctgattctgagatcgagacagagacactttagccataaattgaaagcatctcttcggaaattac	1431
QY	1299	tcgagcagatgagatgctgactcgatgaaagagatctgagagagagacgaagcgaaatagctga	1358
Db	1432	tcgtctgctgtagtcagcagagagcctgtctgtcttcgccacagatgactggaanaacatctgcta	1491
QY	1359	tcgatttatcaatgagattcttcataaattctacatctgtagtagggagaaagagcgctca	1418
Db	1552	tgatttatcatgaaactacactaactaataatttgcacactcgtctgtagagagagagggccca	1611
QY	1419	aatgagtcgagagcgagaaacagagagatcgcgaattgctcgttcgttagtctgaaccocaa	1478
Db	1612	gctgtagctgtagcagaagaaacagagaaatctgcactctcgcggccctggtctgcgaaccocaa	1671
QY	1479	gattctattttagatgaaagctacgctctgccttcgattacaaagacaagatcgactgttca	1538
Db	1672	gattctctctgctgtagtgaagcgaacgctcagctctggaacactgaaagtgaagcagctggttca	1731
QY	1539	agctgcacgcgagaaagcgagcaaaagctcggaactcgaactcgtctgtagacacaccgacttc	1598
Db	1732	ggtgcgcccttgataaagccgagaaagccggacactccattgtagatgactcatcgtttgttc	1791
QY	1599	tactattctgaagctgagatgttggtctgtgcacctaaagatctggaatgctcgcgggagaaag	1658
Db	1792	tacagctctgaaagccgagatgctcgtctgcttctgtagctgtagatgcatctgtagagaaag	1851
QY	1659	agccacatgcgaaactaatctgcaaaacagcaggtctatatttcaactctgtagtgcacaaga	1718
Db	1852	aaatcatgtagaactcatgaaagagagagggcatcttaactcaaacctgtgcacaatgcagac	1911
QY	1719	tattaaaaaaagcctgtaga-acagatgtagtctcaatgaacatattctctctgaaagaaagacca	1777
Db	1912	aagagagaaatgaaatctgagcttgtagaaatgcccactcgtgtagatccaaagatgaaagttagtc	1971
QY	1778	actcaactctctcgacactcgtctgtaaagaag-----catcaagtcagagcttcatgtaaca	1839
Db	1972	cttggagaaatgctcccaaaagatctcaggtcccgattataaataaagaagatcaactcgcag	2031
QY	1830	ggctcgaggaatccaccaactctaaagagaaagatcctt-----	1866
Db	2032	gagtatcatgacgcacaaagcccaagaacagaaagcttgttacaanaagagagacttgaatga	2091
QY	1867	-----cctgaaggtctctctatataaatttaagtataaacaagcttcgatgtgctctt	1919
Db	2092	gaatgataccccacagctctctctcttcggagagatcttcgaagctcgaaactcaactggaattgcctta	2151
QY	1920	tcgtgctctgtagacacatctgctctgtcttctaaatggaactgtctacacagatatttccat	1979
Db	2152	tttctgtgtctgtatatttttgtctattataaagagggccttcgacaacagacttttccat	2211
QY	1980	catcttctgcaaaatataaccatgctt---ggaataatgtataaacccacattaaagca	2036
Db	2212	aaatatcttcaagaactatagggatcttctacccgagatgtagatctcttgacaataaacaagcaca	2271
QY	2037	tcgagtcgaagaattatctacatgtagatctgcatacttcgggtgtagtctgtctgtacagta	2096
Db	2272	gaatagtagaacaatgtttctgtatgttcttagtcttcgtgaattatcttctttataactt	2331
QY	2097	tttcaatgcagagatattcttaacgcagagcagggaggaattttaacgatgagattaaagca	2156
Db	2332	ttttctccacagagcttcaacatcttgccaagaactctgggagatcctccacataaagcggctctgata	2391

OY	2157	cttgcgcttcaagcactgltatatactgaagatattgcgtgtttgtgtgaaaggaanaacg	2216
OY	2157	cttgcgcttcaagcactgltatatactgaagatattgcgtgtttgtgtgaaaggaanaacg	2216
Db	2392	catgtttcttcagacacactgtctgagaagagatgacagctgtgtttgtatgaccttaaacac	2451
OY	2217	cacgagagcttgacaacaacattatagccatatactagacaacattcaagagacaagc	2276
Db	2452	cactgagagcatttgacaacccgctctgcacatgtatgctggtctcaagttaaagggctacagc	2511
OY	2277	ttccagatttgcgtcttcaacacaaaatgcacaacatgtaggactttcaagtatcatcttc	2336
Db	2512	ttccagcttgcgtcttcaacacaaaatgcacaacatgtaggactttcaagtatcatcttc	2571
OY	2337	ctttatatatagatbgagatagatgcacattcctgattctgagatttgcctcaagtcttgcgt	2336
Db	2572	cttaactatagtgtgagcaattaaacacttttactctttagcaattgtatcccatcatcgtcat	2631
OY	2397	gacaggaattgattgaaacccgacgacgacatgctgattcttgccaacaaagttaagcaagact	2456
Db	2632	agcagagagcttgttgaataatgaanaatgtctgtctggaacaagactgtgaagaagaagagact	2691
OY	2457	taagcaatcttggaanaagataagacaactgaagccttgtggaatactacgtaactatagtcatc	2516
Db	2692	agaaagagctctgaggaagattctgtctacagaagcaatcgacaacattccgaactgtgtctctt	2751
OY	2517	aacaagaggaaaaagccttcgacgaacatgtatgaaagatgcttcagactcaacacagaaa	2576
Db	2752	gactcggagagcagaagatttgaatacactgtatgacacagagtcttgacaagtaaccatacga	2811
OY	2577	tactctgagaagaagcacaagtttatgtgaagctgttatgtatcttcagcacttatata	2636
Db	2812	ctctttagagaagaacacacactcttcgggtctcatcttctatcacccaggaatgata	2871
OY	2637	tttgcacatgcacaagaggtttcgatttgcagtccttaactaacagcttggaagcaatgac	2696
Db	2872	ttttctcatagctgcgctgtccgtgttggtgtgctcactctggtggaagaatgattcaatga	2931
OY	2697	cccgagagcgatgtcatagtttctatctgcaattgcatactgagctatgagccatgcgaaa	2756
Db	2932	ctttcagagatgttctcttgttgattcttcacgattatgcttcttggctgcacatgvcagttggca	2991
OY	2757	aagctcgttttgcgtcttcgaatatctcaaaagcaaatcgggggtctgacatctgtttgc	2816
Db	2992	ggtcagatctcatcttctcttcgactacatgcacaagaacatgacagagccacagtcacat	3051
OY	2817	cttgttggaanaagaaacacaatactagacccgacatcaagaagaggaanaagacagcac	2876
Db	3052	gatactatgaaaaaagccctctgattgtacagctcacagccctcaccgctcccaagcaaatac	3111
OY	2877	atgtgaagaggaatttagagttctcgagaagctctcttctcttctatctcatgctgcgccagatg	2936
Db	3112	gttggaaggaatgtgacatttaagatgagtgctgtgttcaactatcccatctgcagcagacat	3171
OY	2937	tttcaactcctgctgactatccctcacgattatgagcgaggaagaacagatagcatctgtggg	2996
Db	3172	ccccgtctcccaaggggtctgacccctgcaggttgaaagaaagggccagacagctggtccctgtagc	3231
OY	2997	gagcagcgagctgttggaanaagcactctgttcaactcttcgagagactttatgaccccg	3056
Db	3232	tagcagtgagcgtgtggaagaagcagattgtgttcagcctcccaagacgctctctatgacccct	3291
OY	3057	gcaagagacaagctgtcttcttgaatgtgtgatagtcaaaaagaatctgaaatgtacagttgctcgc	3116
Db	3292	ggctcgtgttcagtgtaatttgatgacaagaagataaagacacctgaaatgtccagttgctcgc	3351
OY	3117	tttcccaatatagcaatcgtttccctcaagacctgtgtcttcaacgtgcagcatattgtctgaaa	3176
Db	3352	agcacacactgtgacatcgtgtctcaagagacccactctgtttgactgtgcagatgtccgagaa	3411
OY	3177	catgacctatgtgtacaacacgctgtgtgtgcattagatgagatgacaagaagccgcaaa	3236
Db	3412	catgtccctatgtgaaacaacacgctgggtctgtatcatactgaagagattatgtcagagcagccaa	3471
OY	3237	tgcagcaaatatccatctcttttatgtgaaggtctcccttggaatactcaacacacaagttgcg	3296



```

Db 3472 ggaagccaaacatacaccactcaccagacactccctcgagaaatacacaacacagtagag 3531
Oy 3297 actgaagagacacagcttctgagccagaaacagagtagctatcttcaagagctct 3356
Db 3332 agacaagaagacacagctctctgagccagaaacagagtagctatcttcaagagctct 3531
Oy 3357 tctccaaaaccccaaaatttatctgtgagtgagccactcagccctcagataagcag 3416
Db 3392 tgttagacagctcattatctgtcttgatgaagctacatcagctctgtatagaag 3631
Oy 3417 tgaagaagtggttcagatgcctctgataaagccagagcgggaagcagatgctatgct 3436
Db 3652 tgaaaagtggttcacaaagacccctgacaaagccagagagcgcgcactgctatgctat 3711
Oy 3477 cactcagagctctctgctcagatcagagcagagattgatgctgtcttcgacaatgaaa 3536
Db 3712 cgcaccacgcttgcacacacacagagatgcagattatagtggttctcagaatgcaa 3771
Oy 3537 gataaaggaacagagacatcacaagagctcctgagaaatcagacatataatttaagt 3596
Db 3772 agtcaagagagcagatgcacacacacagctgctgcccagaaagcattatcttccat 3831
Oy 3597 agtgaatgcacagtcag 3613
Db 3832 ggtcagtgctcagagctg 3848

RESULT 5
AAD03489
ID AAD03489 standard; cDNA; 4317 BP.
AC
XX AAD03489;
XX
DT 13-JUN-2001 (first entry)
DE Dog P-glycoprotein (PGP) cDNA #2.
XX
XX Dog P-glycoprotein; PGP; multidrug transporter; MDR1;
KM Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
XX drug bioavailability; transgenic animal; genetic model; ss.
OS Canis familiaris.
XX
XX Key Location/Qualifiers
FH 70..3912
FT CDS /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #2"
XX
XX WO200123540-A2.
XX
XX PD 05-APR-2001.
XX
XX PF 28-SEP-2000; 2000MO-US26767.
XX
XX PR 28-SEP-1999; 9905-0156510.
XX
XX PA (GENT-) GENTEST CORP.
XX
XX PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Paten CJ;
XX WPI; 2001-235373/24.
XX DR P-PSDB; AAE00304.
XX
XX PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors -
XX
XX PS Claim 1; Page 66-72; 111pp; English.
XX
XX CC The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX CC The invention also includes fragments and biologically functional
XX variants of dog P-glycoprotein. PGP and their nucleic acids are

```

```

CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC "knockout" phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

```

```

Query Match 36.5%; Score 1320.6; DB 22; Length 4317;
Best Local Similarity 62.5%; Pred. No. 0;
Matches 2187; Conservative 0; Mismatches 1259; Indels 51; Gaps 6;

```

```

Oy 162 caggttgacccctgattatgttggaatagtggtgctgcttgatttggatcacaca 221
Db 408 caggttgacccctgattatgttggaatagtggtgctgcttgatttggatcacaca 467
Oy 222 gattcctgtgataatactgcagcagcagacagacagagagattcgaagaacagtttt 281
Db 468 gtttcaatctgtgctgcagcagcagagagacagacatcacaatcgaagaacatttt 527
Oy 282 tcaatcagtttgcagacagacatcagctgttgatagctgagacatcggtgaactaa 341
Db 528 tcaatcagtttgcagacagacatcagctgttgatagctgagacatcggtgaactaa 587
Oy 342 cactcagatgcacacattg---acaaatcagtgatgattggaataagattgctt 398
Db 588 caccggtccacagacagatgctcccaaaacaaagagagatgagcacaagttggaat 647
Oy 399 gtgttccaaacatgctcatttctcagatggcctgcagctgttggttggatgaaggctg 458
Db 648 gtcttccatcacaatagcaacatttccacggttataagtggtggttacactgtgtg 707
Oy 459 gaaactacccctagctcattccacagctctccttataatagctcagcagcagatg 518
Db 708 gaaactacccctagctcatttctcagatggcctgcagctgttggttggatgaaggctg 767
Oy 519 tttagagtgctatcctcattgacagtaagaaatagtgctcattccaaagcttgaggc 578
Db 768 ggcagaagatcattcattcatttgcacacagacccctgttggaattcgaacccgcatgt 827
Oy 579 tgtgcagaagaagctctgtcatcaatcgaacagctatagcctttagggccagagaa 638
Db 828 agtagctgaagaagctcttagcagcaatcgaacagctgattgctccttgtaggacaagaagaa 887
Oy 639 agaactccaagatatacagaagatcctcaagaatgcaagaagatttggcataaaagagac 698
Db 888 agaactccaagatatacagaagatcctcaagaatgcaagaagatttggcataaaagagac 947
Oy 699 tatagctcaaaaagtgctctgttggtgctgtaacttcttaataatggaactaagtagact 758
Db 948 tatagctcaaaaagtgctctgttggtgctgtaacttcttaataatggaactaagtagact 1007
Oy 759 tgccttggatgaagacccctgattccttaagagagacccctgataccatcagagac 818
Db 1008 ggccttctgtatagagacccctgctgctcctcag-----tgaatactatgtgaca 1061
Oy 819 tgtctctgctgttctctttagtgtaatcacaagcattatgcatgagcagcagctcc 878
Db 1062 gttactcactgtctcttcttcttgaatgaatgagcctttagttagcagcagctcc 1121
Oy 879 tcaattgaaacctgcagatagccgagagagcgccttcatatttccaggtatga 938
Db 1122 aagcatggaagcattgcagaagcagaagagagcagcttcatgaaacttcaagaataattga 1181

```

QY	939	taagaaaccagatatagataaactttccaagcttggagatataaacttgataccaatgaag	998
Db	1182	caataaaccagacattgacagactatcgtgaaggttgacataaaccagataatactaaagg	1241
QY	999	aactgtggaatttaaaaatggtttcttcaattatccataccaagaccatctacagattcc	1058
Db	1242	aaatttggaaatccaanaaaatglttaacttccagtttaacctcttcgnaaaagatlaagactt	1301
QY	1059	gaaaggtctaaactccaagataaactcttgagagaaacagctgcgcttgcgtctcaatgg	1118
Db	1302	aaagagctcaaccactgaaggttcagaggttggagagacagttgcgctggttggaaacagtg	1361
QY	1119	caatgtggaagagtaagtaagcttccagctcttcagagaggtataatgataccgatagtgctt	1178
Db	1362	ctgcggaagaagacagacaccgttgaagcttgaagaggtcttataatgaccccaagatggcat	1421
QY	1179	tatcatgtgtgataggaatggaactcaagagcttaataatgtcggcattatcgagacatat	1238
Db	1422	ggtctgatatcgagacagagacactatgagacataatgaaagacacttcgggaattac	1481
QY	1239	tgagtggtttagtcaagagagcctgtttcttgcggacccacatcagatlaacaatatacagta	1298
Db	1482	tggtgtgtgtgagtcagagagcgtgtgtgttcttcacacagataagcttgnaaacattcgcta	1541
QY	1299	tggagcagatgatgtgactgataagagaatggaagagcagcaaggaagacaaatgcgtata	1358
Db	1542	tgcgcgcgaaatgtcacaccatgtgatatgatatgagaagctgtttaaagagcaatgctota	1601
QY	1359	tgatttatatactgagtttccctataataattatatacatcttgtagagggaaaaagagctca	1418
Db	1602	tgaatttatatacgaactacactcaataaatttgcacactcgtgttgagaagaaagggcccg	1661
QY	1419	aatgataggagggcagaanaacagaggaatgcgaatttgcgttcgttagtctgaacacccaa	1478
Db	1662	gctgagtggtgagcagcaaaaagagaaatgcgcaattgtcgtgcgcccgtgtgcgaacccaa	1721
QY	1479	gattctgattttagatgaaggtctacgcttcgccttgattcagaagaacagtcagctgttca	1538
Db	1722	gattctcttcgtgatgtagagcaacgctcagctcttgcacactgaagtgaagcagtgattca	1781
QY	1539	agctgcaccttggagaagggcgaagcaaggtctcgagctacaactcgtgtgtagacacccgacttc	1588
Db	1782	gggtgcccgtgataaagccgaanaaaggccggacataccatttggatagctatcgttttc	1841
QY	1599	tacattccagaagtgcagatttgaatttgcacccaagaagttgaatgctgcgagcggaaggg	1658
Db	1842	tacagttcgtcaatgcccgaatgtcattgtcgtgttttgaatgtagagtcatttgcggagaag	1901
QY	1659	agcacatgctgaaactaatgtgcgaacaaacgaaggtctatataatcacttgtgatgcagga	1718
Db	1902	aaatcatgatgaactctatgaaagaagggcatttaacttcaaaccttgcacaaatgcagac	1961
QY	1719	tattaaaaagcgtataga-acagatgtagtcaatgatcatatctctatgaaagaagacca	1777
Db	1962	aagsgagaaatgaaatgtgagttagaanaaalcgcacttgtgataccaagaatgaaagtgaagc	2021
QY	1778	actcaactccctgcgaactcgttgaagag-----catcaagtaagaacttcatatgacaa	1839
Db	2022	cttggaaatgtctccaanaaagaatccaggttccaggttcaataaanaagaatcaactcgcag	2081
QY	1830	ggtctgaggaatlcacccaatcaaaagata-----	1860
Db	2082	gagtatcatatgacacccaagggccaagacagaagccttgttacaagaaggacttgaatga	2141
QY	1861	-agcttccctgaagctctctatataaanaattttaaagttaaacaagcctgaaatgctt	1919
Db	2142	gaatgtctctcagatttctcctcttcggagagatctcgaagcgtgaacactcaatgaaatgctta	2201
QY	1920	tgtgtctctgggagcatgtgctctcttctcaaatggaactgtctacgaattatctcat	1979
Db	2202	tttctgtgtgtgtatatatttctgtctattataaaccggaagcctgcgaacagcatttccat	2261
QY	1980	catcttgcgaanaaattlaaacagtglt---ggaaataatgataaaccacattaaagca	2036

Db	2262	aatttttccaaaggaattataaggaattctttcccgagatgaggaattccctgaacaaacgaca	2321
Oy	2037	tgaagcagaatattatccatgatactgcatctttgggtatgttctgttcgagtia	2096
Db	2332	gaatagataacatgtttctgtatgtttcttccagtccttgaaattatcttttataact	2381
Oy	2097	tttcatgcaaggatattattacgagcagagcagggnaattttaacgatgatataagaca	2156
Db	2382	tttccctccaggagcttccatttggccaagctggggagatccctcaataagcggctctgta	2441
Oy	2157	cttgccttcaagccatgttatatacagatattgtccgtgtttgtgtgnaaaggnaaacag	2216
Db	2442	catggttttccaagatccatgtctgagacaagatgccaagctgtttgtatgaccttaaaac	2501
Oy	2217	caacgagagctgtgacaacaatatacagctagaatagacacaatattcaaggagacaacg	2276
Db	2502	caatggagcatatgacaacccgctgtgcacaatgatgagggctcaagttaaaaggccataag	2561
Oy	2277	ttccagaattggcgtcttaacacaaaatgcacaatcaatgaggaacttcaagtatcaattc	2336
Db	2562	ttccagagcttgcgtcatcttaccagaatataagaaatcttggagacagcatattatatac	2621
Oy	2337	cttatataatgatatgggaagtatgcattccctgattctcgaattgttcagtaatttgcgcgt	2396
Db	2632	cttaatcatatgtgtggaattaaactttacactttacactttgaagacttgaacatcatatgcaat	2681
Oy	2397	gacaggaatattgnaaacccgacgacaatgatgattgtccacaacaagaataagacaagact	2456
Db	2682	agccagagatctgttgaataatgaaatgtttgccttgagacaagcaccgtgaagaataagaaagact	2741
Oy	2457	taagcatatcttgaanaagataagacaacttgaaagcttttgagaaatatacgttactatagtcat	2516
Db	2742	agaagagagcttgaggagaagtttctacagaagccatgcgaacattccgaactgtgtttctt	2801
Oy	2517	aacaagggaaaaagccttcgaagcaaatgtatgtaagaagatgtcttcagactcaacacagaagaa	2576
Db	2802	gacttcggagagcagaagatttgaatacattgtatgcacagagtttgcagaattacatacacaagaa	2861
Oy	2577	tactctgaagaagaacacagaatttatgtgaagcgtttatgattccagcagcatccattatata	2636
Db	2862	ctctcttgaggaaaacacacacattcctgggggtctcaatttctatcaaccaggccaatgatata	2921
Oy	2637	ttttgcctatgcagcagaaggtttcgaatttggagccatattaattcaagctgagacgaatgac	2696
Db	2922	tttttccatgtcgtgcgtttccggtttgttgctgactattgttggccaatgagttcatgaa	2981
Oy	2697	ccccagagggcatatgcatatttttactgcgaattgcataatgagactatgagccatctggaaa	2756
Db	2982	ctttccagagatgtctcttttgttatcttccagctatgtcttcttggccatgagcagtgaggaca	3041
Oy	2757	aagcgtcttcttgcctctctgaatactccaaagccaatccggggcgctgcgcagctgtttgc	2816
Db	3042	ggtcagatcatatttgcctccctgactatgccaagaagccaagaatatacgaagcccaagtcataat	3101
Oy	2817	cttcttggnaaaagaaaccaaataatagacagccgacgtcaagaagggnaaaagccagacac	2876
Db	3102	gatcatattgnaaaaaagccctctgtatttgcacagctacagccctcaagcctccaacaataac	3161
Oy	2877	atgtgaagaagaaattagaaatttcgaaagctcttctctctatccaatgtctgcccaagatgt	2936
Db	3162	gttggaaagaaatgtgcacatttaatgatgagtcggttccaacttcccaatcgcagcagacat	3221
Oy	2937	tttatactccgtgtgcattaccctcagtaatttgagcaggaagaagaagatgagtaagtttggg	2996
Db	3222	ccccgtgtctccaagggtcggagcctcgaggttgaaagaaagggccaagacgctggccctcgtaag	3281
Oy	2997	gagcagcgcgctgttggaagaagcaccttctgtccaactcttcgacagagacttatagcacccgt	3056
Db	3282	tagcaggtgagcgttgaggagaagcaagttgttcaagctccctagaagccttctatagaccctt	3341
Oy	3057	gcaagacaaagtctgttctgtatgtgtgtgatagcacaaagaattgaagtacagttgctcgg	3116



Db 1115 agtactcactgtlatcttctctgtlatiaattgaggcctttagtctgtgacagacatctcc 1174  
Qy 879 tcaacttgaacacctgcgaataagcccagaggcctcttcaattattccaggttatga 938  
Db 1175 aagccttgaagcaatttgcgaataagcagaggcagcttatgaactcttcaagaatga 1234  
Qy 939 taagaacaacccagatagaacttcttccacagcttgatataaaccttgatccatagaag 998  
Db 1235 taataagccaagtatctgacagctatctgaagagtgaggcaacaacagataatataaggg 1294  
Qy 999 aacgttggaattttaaataatgttcttcttcaatlatccatcaagacatcatatcaattcc 1058  
Db 1295 aaatttggaaattcagaataatgttcaacttcaacttaccatctcgaataaagaagttaagctct 1354  
Qy 1059 gaaagctctgaatctcagaatlaagtctcgagagacagctgcgtctgctgcgttcaatg 1118  
Db 1355 gaaaggccctgaaacctggaaggctgacagctgaggcagcgttgcccggttctggaataacgtg 1414  
Qy 1119 cagctggagaagtaacgctgacagctctctgcagagcttatatgaaccgatatgagctt 1178  
Db 1415 ctgtggagagagcacaacacagctcagcttgatgacagagctctatgacccacagagggat 1474  
Qy 1179 tatcatgttgatgagaaatgacatcagagctttaaagtgccgacatctatcgagacatcat 1238  
Db 1475 gctcagctgttgatgagacagatataaggacataaagttaagcttcttctacgggaatcat 1534  
Qy 1239 tggagctgttgatcgaagagcctgttctgttctggggacacacatcaatataatagta 1298  
Db 1535 tgggtgtgtgtgacgagacacactgtatgttctgcccacacagatagcttgaataactctgcta 1594  
Qy 1299 tggacgaatgatgacgtgacgtatgaaagatgagagagcagaagcagaagcagaatgctta 1358  
Db 1595 tggcgcgtgaaataatgacacatgagatgagatgagagaagctgttcaaggagagccaaatgctta 1654  
Qy 1359 tgaatttatcatgtgagtttctcctaataacttaatacatgtgtaaggaggaataagagctca 1418  
Db 1655 tgaatttatcatgaaactgcacctcataaatttgacacacctgttggagagagagcgccca 1714  
Qy 1419 aatgagctgagagcagaacacagagatcgcaattgctctgcttgaagtctgaaacccca 1478  
Db 1715 gttgagctgtgtgagcagaagcagagatcgccatctgacgtgcctgttgcgaacccca 1774  
Qy 1479 gatctgatttatgagtgagctacgtctccctcgatccgaataagcagctgcgttca 1538  
Db 1775 gatcctcctgtctgagtgagcgcaacgtctcagctctgacacagaacgagcagctgttca 1834  
Qy 1539 agctgacactggaagagcgagcaaaagctcggactaacaatcgtgtgtagcacacagacttc 1598  
Db 1835 ggtgctctggaataagccagaaagctcggaacacacatgtgtagctcatcgtttgtc 1894  
Qy 1599 tactatttgaagtgagatttgaatgtgaaccttaagagatggaatgctggtcggaagag 1658  
Db 1895 tactgcttgaatgtgagctgcatcgtctggttctgatgtgagtcattgttgaagaaag 1954  
Qy 1659 agcacatctgaacacatgagcaaaagctcatatattatcaactgtgtgtatcgacga 1718  
Db 1955 aaatcatgatagaactcagaaagagaaagacattacttcaaaactgttgcacaatgacgac 2014  
Qy 1719 taataaaaacgtgataagacagatgagtgacatgacatactta----- 1762  
Db 2015 agcagaataatgaagtgaatttagaataatgacgtgataatccaaaagtgaattgatgc 2074  
Qy 1763 -ctgaagaagaagacaaactacactcctctgcactctgtggaagacata----- 1810  
Db 2075 ctltggaatgttcttaaatgtatcagaatlcagatctaaataagaaataaagatcaactcgtag 2134  
Qy 1811 -----agtacagcttcatatgacaagctgaggaatccaccacatcaagaagat 1859  
Db 2135 gagtctcgttgatcacacagcccaagacagaagaagcttagtaccaaaagggctctgtatga 2194  
Qy 1860 aagctctccctgaagctctctctatataaaaatttaaagttaaacaagcctgaatggcctt 1919  
Db 2195 aagtataccctcagcttctccttcttggagagatatgaagcctaattcaactgaatgacctta 2254

Qy 1920 tgtgttctcgggagacattgagctctgtcttaaatgaaactgttaccagatatttccat 1979  
Db 2255 ttgtgtctgtgtatatttgttgcattataaaygaagcctgtgacaacagacttgcacat 2314  
Qy 1980 catcttgcgaataaataaaccatgtt--ggaataatgataaacaacatlaaagca 2036  
Db 2315 aatatttccaaagatataagggattttaaagaatgatgtctcctgtaacaaacagca 2374  
Qy 2037 tgaatgagaataattatccatgatatctgtcatatttgggtgtatattgtctgtcaagta 2096  
Db 2375 gaattgaactgttcttactatgttctcagcccttgaattacttcttatttaccat 2434  
Qy 2097 ttctcagagagatattttaagcagagcaggggaaattttaaagcagatgaagca 2156  
Db 2435 ttctctcagaggttccacatlttggaacagctggaagatccctcacaagcggctccgata 2494  
Qy 2157 ctgtgcttcaaagacatgttatatacagagatatcgcctgttgaatgaaagaaacag 2216  
Db 2495 catggttctcgatccatgactcagacagagatgtgagttgtgttgaaccttaaaacac 2554  
Qy 2217 caaaggaagcttgacaacaataatagccatagatatagcaacaattcaaggaagcaag 2276  
Db 2555 cactggaacatgactacacagctcgcgaatgatgtctcaagttaaagggtctatag 2614  
Qy 2277 ttccagagcttgagctcttaacaaatgcaactacatgagacttcaatcatcttc 2336  
Db 2615 ttccagctctgtgttaataccagaaatagcaaatcttggagacagaaatataatc 2674  
Qy 2337 cttaataatgatatggagatgacatctcctgacttctgagatgttgcctgaacttgcgt 2396  
Db 2675 cttaactatgttctgacacttaacactgttactctcttaagcaattgacacatcatcgaat 2734  
Qy 2397 gacaggaatgatgaaaccgcagcaatgactgagatttgcacaagaataagcaagaact 2456  
Db 2735 agcagaagcttgtgaataagaaatgttctgtgaacaacactgtaaaagataagaaagact 2794  
Qy 2457 taagcatgctggaagatagacacactggaagcttggagaataatagctactatgttcat 2516  
Db 2795 agaaagctctgggaagatcgcacacagaaacaaataaactctcgaaacgttgttctt 2854  
Qy 2517 aacaagggaataaagccttcgagcaaatgtatgaagaatgtcttcagactcaacacagaa 2576  
Db 2855 gactcaggaagcaagaatttgaacatagtatgtctgaaggttgcaggtgaacatacagaa 2914  
Qy 2577 taactcgaagaagcagacagatataatggaagctgttaatgcatcagacatgtcttata 2636  
Db 2915 ctcttggagaaagacacatctcttgaaattacatttctcctcacccagcgaatgtga 2974  
Qy 2637 ttgttcataatgacagaggttctgatttggagcctatttaactcaagcttgagcgaatgac 2696  
Db 2975 ttcttccatagtgtgaatgttccggttggagcactctgtgagcaataacatcaatgag 3034  
Qy 2697 cccaagagcagctcatagtttcttaacgcgaatgcatatgagagctatgagccatcgaa 2756  
Db 3035 ctttgagagatgtctcgttaagatatttcaagctgttcttgggtgcacagccgtgggca 3094  
Qy 2757 aacgctgttctgctcgtgaatatctccaagccaatccgggggtcgtgcacatgttgc 2816  
Db 3095 agtcaatcttctgtcctgactatgccaagccaataatcaagcagccacatcatcat 3154  
Qy 2817 ctgtgttggaagaagaacaaatataagacagccgacgttcaagaaggaagaagccagac 2876  
Db 3155 gatcatgtgaataaacccttgaatggaagctcagacgcggaagcctcaatgtcgaacac 3214  
Qy 2877 atgtgaaggaatttagagtttgcgaagctccttctctctatccatltgcgccagatc 2936  
Db 3215 attggaaggaataatgacaaatttgtaaggttgaatttcaactatccacccagcaggaat 3274  
Qy 2937 tttaactcctggtgcttatccctcagatattgagcgaaggaagacagatgacttgtgg 2996  
Db 3275 cccaatgtctcaggaactcagccttggaaggtgaagaagggccagacgctgtcgtgtgg 3334

[illegible]

Db 1240 agaacttgaaagtgacaacaaatttagaagaagctaaagaattgggataagaagc 1299  
QY 699 tatgcttcaaaagtgctctctgtgctgtgctgcttctttagaattgaactgaact 758  
Db 1300 tattacaaccataattctataatagtgctgcttctctgtgactatgatactatgacct 1359  
QY 759 tgccttttgatggaacctcttgatcttaattgagaacctgatatcatcagggac 818  
Db 1360 ggcctctcgtatgagcacacctgtgctctcctaaggga-----atatctatgga 1413  
QY 819 tgttctcgttcttctttagtgtaatccatagacagcttatgtacattggagcagctccc 878  
Db 1414 agtaactacgtatcttcttctcgtatattagggtctttagtggttgacagagctccc 1473  
QY 879 tcaacttgaaccttcgacaatagcccgagagctgtccttcaatattccaggtatuga 938  
Db 1474 aagccttgaagcatttgcaaatgcaagagagcagcttatgaaacttccaagataatga 1533  
QY 939 taagaagaccggtatagaatttccacagcttgatataaacttgataccatagaag 998  
Db 1534 taataagccaagctatgcagactatcgaagaagtgggccaacacgaalaaatlaagg 1593  
QY 999 aactgtgaatttaaaaaagtcttcttcaatatacatcaagaacatcatcaagatct 1058  
Db 1594 aaatttggaattcagaatagtcttacttcaagttaccatctcgaaaagaagtaagatct 1653  
QY 1059 gaaaggtcgaatcgcgaatagatctgagagagaacagtcgcttggtcgtctcaatgg 1118  
Db 1654 gaaggccctgaaacctgaaagtgcgaagtgagagacggtgagccggtgtgaaacagtgg 1713  
QY 1119 cagtgaggaaggtagcgtgagtcagactcttcgagaaggtatataatgctcgatgagct 1178  
Db 1714 ctgtgggaagacaaacaaagctccagctgcgaagagctctatgaccccaagaggagct 1773  
QY 1179 tatcatggtgagtgagaaatgacatcagagctttaaattgvcgcaatlatcgagacatat 1238  
Db 1774 ggtcagtggtgagcagagatataggaacataatgaatgaagttcttcaacggaatact 1833  
QY 1239 tggagtggttagtcaagagcgtgttctgttcggagccacacatcagtaacaatatcaagt 1298  
Db 1834 tgggtgtgtgagtcaggaacctgtatgttctgccaacgaagatagctgaaacatctcgta 1893  
QY 1299 tggagagatgagtgcactgataagagatgagagagagagaggaaggaacatgctga 1358  
Db 1894 tggcgtgaaaaatgcaacatgagatgagaaagctgccaagggccaatgacctga 1953  
QY 1359 tgactttatcagtaggttctccataaatttatatacattgtagggaaaagagagctca 1418  
Db 1954 tgactttatcagaaactgcctccataaatttgaacacctggttggtagagagaggggacca 2013  
QY 1419 aatgagtgagggcagaacaaagagatgcgaattgctcgttgccttagtgcgaaccccaa 1478  
Db 2014 gttgagtggtgagcagaagcagaagatgcacattgcaagctgcccgtgttcgacaacccaa 2073  
QY 1479 gattctgatttagtgaggtcagctcgtcccttgatctcaagaagaagtcagctgttca 1538  
Db 2074 gatctctcgtcgtgagtgagggccaagctcagcttggacacagaagcagtgagttga 2133  
QY 1539 agctcactgagaaagcgaagcgaaggtcgaactcaatcgtgtagcagacacagcttc 1598  
Db 2134 ggtgctctcgtgataagcagaagaaggtcggacacatgtgatagcctcacgtttgtgc 2193  
QY 1599 taactctcgaagtgcagatttgatctggaacccctaaagagatggaatgctgvcggaagaag 1658  
Db 2194 taacagtctgaatgctgagcatcgtgttctgatagtatgagatcattgttgaagaag 2253  
QY 1659 agcacatctgagactaaatgagcaaacaggtctatatattctacttgatgctcagga 1718  
Db 2254 aaatcatgatagacatgaaagaaagaaagcatttaactcaaatgttgcacaaatgcaagc 2313  
QY 1719 taataaaaagctgataagacagatgagtcacatgacatactta----- 1762  
Db 2314 agcagaagaatgagatgaaattgaaatgcaagctgataatcccaaaagtgaaattgagtc 2373

---

QY 1763 -ctgaagaagaagccaactccttccctcgtcagctctgtggaagagcatca----- 1810  
Db 2374 ctgtggaatgcttccataaagatccaagatccagctcaataaagaataaagatcaactcgtag 2433  
QY 1811 -----agtcagacttcatgacaagcgtgaggaatccaccacataaagaat 1859  
Db 2434 gagtgcgtgtagtacaagcccaagacgaagaagcttagtaccagaagagccttgatga 2493  
QY 1860 aagttccctgaagctcctctatataaaatttaagttaaacgaagcctgaaatgcctt 1919  
Db 2494 aagtaactccagttcccttlttgagagattatgaagcctaataattcaactgaatgcctta 2553  
QY 1920 tgtgttctcgggaacttggtcgtctcgttcaatggaactgttccatccagttttccat 1979  
Db 2554 ttgtgtgtgtgtatattgtgcacattataaattggaagcccgcaacagcatgtgcaat 2613  
QY 1980 catcttgcanaaatataacacatgltt--ggaataatgataaaccacataaagca 2036  
Db 2614 aatatcttcaaaagatataggggttttacaagaattgtagtccctgaaacaaagcaga 2673  
QY 2037 tgaatgcagaaatltatccatgatatctgcataattvgggttaltgtcttgcaagta 2096  
Db 2674 gaaatgaactgttctcaacttgttccagccttggaattattcttattatcaact 2733  
QY 2097 ttcatgcagagattatatttaagcagagcagggaaatttaacagatgagattaga 2156  
Db 2734 ttctctcaaggttttccaatlttgcaaaagctggaagacccacaaagcgtcccgata 2793  
QY 2157 ctgtgccttcaagcaltgtlatcatcagatatctgcgtgttgaatgaaagaagaaacag 2216  
Db 2794 catgttlttcgatccatcagctcagacagagatggtgtttagttagacccttaaaacac 2853  
QY 2217 cacagagagcttgacaacaaatagccacatagatatagcaataatccaagagcaag 2276  
Db 2854 cactgagagatgactaacacagcgtccgaatgatgctgccaagtaagaaagggtcagag 2913  
QY 2277 ttccagagatgctgcttcaacacaaatgcaactaacacagggacttccgtatcaattc 2336  
Db 2914 ttccagagcttgggttaattaccacgaatatagcaaatcttggagacaggaattatattc 2973  
QY 2337 cttaataatgagatggagatgacatctcgtatctcagatgttcgcagacttgcgt 2396  
Db 2974 cttaactctgtgttggcaactaacatcgttactcttagcaaatgtgataccatcattgcaat 3033  
QY 2397 gacaggaatgattgaaacccgagcaatgagcttgatgtgcacaaagaagaagcaagact 2456  
Db 3034 agcagagatgttgaatgaatgaatgaatgtgtctgacaaagcactgaaagataagaagaact 3093  
QY 2457 taagcatgctgaaagaatagcaactgaaagccttggagaatatagctactatgctcat 2516  
Db 3094 agaaggtgtgtggaaagatgcgtactgaagcaatagaanaaacttccgaacggttctt 3153  
QY 2517 aagcaagggaaaaagccttcgagcaaatgfatagaagagcttcaagaccacaacaga 2576  
Db 3154 gactcaagagcagaagttgaaacatatgatatgctcagaatttgcagatccatatacaga 3213  
QY 2577 taactcgaagaagcacaagatatltggaagctgttatgtgatacgaagccttatata 2636  
Db 3214 cccttggagaagaagcacaacatcttggaaattacatttcttccataccaggaatgtgta 3273  
QY 2637 ttgttcctatgcagcagaggttccgatattgagagcttatcaatcaagcctgagcaatgac 2696  
Db 3274 ttlttccctatgtagtgttccgggttltggaagcctactcgtgtgacataaactagag 3333  
QY 2697 cccaagagggcatgtcatagttttactcgaatttgataatgtagatgagctatgccaatgga 2756  
Db 3334 ctttgaagatgtctcgtatgattatcagactgttgccttgggtccatgvcggtgggga 3393  
QY 2757 aacgctcgttlttgctcgtgaatattccaaagcnaaatcgggggctgctgctgttgc 2816  
Db 3394 agtcagttcaattgtcctcgtcactatgccaagcnaaatatacagcagcccaatcatcat 3453

```
OY 2817 cttgttggaaagaaacccaatatagacagccgacgcaagaagggaagaaacccagcac 2876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3454 gatcattgaaacaaaccccttgattgacagctacagagcgaagcctaagccgaacac 3513
OY 2877 atgtgaagggaatttagagtttgagaagctcttcttctatccagtgcgcagagt 2936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3514 attggaaggaaatgacatttgtagaagttgtatttaactatccaccgacggagacat 3573
OY 2937 ttcatccctcgtggttactccctcaglatgagcgaggaagaacaglaagcattggtg 2996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3574 cccactgcttcagggactggtgctgaggtggaagaagggcgcagacgtgctcgtggtg 3633
OY 2997 gagcagcgtctgtggaagaagaacactctgtcaactctcgcagaagacttaagccccc 3056
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3634 cagcagtgctgtggaagaagacagtgccagctcctcgtgagcggttctacgccccct 3693
OY 3057 gcaagagcaagtgctgttgaatgtgtgtagtgcgaagaagattgaattacagtgctcg 3116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3694 ggcagggaaagtgctgtctgtgacgaagaataaacgcgactgtagtctcagtgctcg 3753
OY 3117 ttcccaatagcaatcgcttctcacaagacgtgtgctcttcaactgcagcattgtgaa 3176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3754 agcacactggtgactcgctgtccagggagccacactctgttgcagcagcattgtgaa 3813
OY 3177 catgcctatgtgtgacaacagcggtgtgtgacattagatgagatacaagaagccgcaa 3236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3814 cattctatgagagacacacagccggtgtgtgtacaggaagagatcgltgaagcgagaaa 3873
OY 3237 tgcgcgaatatcatctcttatttgaaggtctccctgaggaataacacacagaattgg 3296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3874 ggaagccaaactacatgctcttcaacgagctacgtccctaataatagcactaaagtag 3933
OY 3297 actgaaggagacagcttctcgcgcgcagaaacaaagactaatttgcaagggtct 3356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3934 agacaaaggagactcagctctctgtgtgcacagaacaaagcgttcataagctgtgcct 3393
OY 3357 tctcacaacccaatatttattgtgtgagtgagccacttcagccctgataatgacag 3416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3994 tgttagacagctcatatttctgtctttagtagaagccagctcgtgatacagaag 4053
OY 3417 tgaagaagtggttcagatcgtgccttgataagcgagcggaaggaatgctcctagtg 3476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4054 tgaagaagtggttcagatcgtgccttgataagcgagcggaaggaatgctcctagtg 4113
OY 3477 cactcacaaggtctctcgtacatcagacgagcagattgtagtggttctgcacaatgaa 3536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4114 tgcacacgcctgtccaccatccagaaatgcagacttaagtgtgttcagaatgacag 4173
OY 3537 gataaaggaaacagaatcatcacaagctcctcgaagaatcgaagacataatttaagt 3596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4174 agtcaaggagatgtagcagcatcagcagctgtgtgcacagaagaagcattctttca 4233
OY 3597 agtgaatgcagatcag 3613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4234 gttcagtgccaggtcg 4250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
AA294738
ID AA294738 standard; CDNA; 4646 BP.
XX
XX AA294738:
XX
XX 01-AUG-2000 (first entry)
XX
XX Human ATP binding cassette ABCB1 (MDR1) cDNA.
XX
XX ABCB1: ATP binding cassette; human; cholesterol; lipid disorder;
XX atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
XX lupus erythematosus; diagnosis; gene therapy; MDR1;
XX multidrug resistance; chromosome 7q21; ss.
XX
OS Homo sapiens.
```

```
XX XX
PN MO200018912-A2.
XX
XX 06-APR-2000.
XX
XX 21-SEP-1999; 99MO-EP06991.
XX
XX 25-SEP-1998; 98US-0101706.
XX
XX (FARB ) BAYER AG.
XX
XX Schmitz G, Klucken J;
XX
XX WPI; 2000-293151/25.
XX
XX Adenosine triphosphate binding proteins useful for identifying agents
XX PT for treating atherosclerosis and other inflammatory disorders -
XX
XX Claim 9; Page 110-112; 154pp; English.
XX
XX PS
XX
XX CC The present sequence is that of human ATP binding cassette
XX CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a
XX CC differential display method in which monocytes from peripheral
XX CC blood were subjected to macrophage differentiation and cholesterol
XX CC loading with acetylated low density lipoproteins and subsequent
XX CC degrading with high density lipoprotein (HDL3) to identify
XX CC cholesterol sensitive genes. The gene maps to chromosome 7q21
XX CC and is also termed MDR1 (multidrug resistance). The invention
XX CC provides cholesterol-sensitive ABC genes (see AA294734-63). These
XX CC genes, and polypeptides encoded by them, can be used for diagnostic
XX CC assays to screen for pharmacologically active modulator compounds
XX CC useful for the treatment of lipid disorders, atherosclerosis or
XX CC other inflammatory diseases such as psoriasis and lupus
XX CC erythematosus.
XX
XX SQ Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;
XX
XX
XX Query Match 36.3%; Score 1316.2; DB 21; Length 4646;
XX Best Local Similarity 62.3%; Pred. No. 0;
XX Matches 2178; Conservative 0; Mismatches 1268; Indels 51; Gaps 5;
OY 162 caggtgacctgtatattgttgaatagtggtgctgctgttttggtaacata 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 760 caggtatgctcattatcagtggaatgtgtcggggtgctgtgtgtctacatca 819
OY 222 gattccctgtgatttaactgcagcagacagacccaagagattcgaacagtttt 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 820 gtttcaatttgggtgctgcagcgtggagacaaataacaaattgaacaaagtttt 879
OY 282 tcaatcagtttggacagacatcgctgtgtgatagcgtgtacatcgtgaactaa 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 880 tcatgctaatgacagagagatagctgtgtttagtgcagatgttggggagctaa 939
OY 342 cactgcagatgacagatgac---aaatcagtgatgttggagtaattgtcct 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 940 caccgacttacaagatgtgtcttaagatcaagttatgtgtgcaaaattggat 999
OY 399 gttgttcaaaacatgctacttcttgcattggcctgcagcattggttgggaaggctg 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1000 gtcttcaagcaatgcaaatlttcaactgggttttaagtaggatttaacagtggtg 1059
OY 459 gaaactcacctagtagctatccacagctctccttataatgtgctcagggacagatg 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1060 gaagtaaccccttgatgtttggcaccagctccttcttgactgcagctgctgtctg 1119
OY 519 ttctagatgtcatcattgacagtaaggaatgaatgctctatccaaagctggggc 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1120 ggcagaagctactatcttaactgaataagaccttcaagcttgaagaagcgtagg 1179
OY 579 tgtgcaagaagaagctgtgtcatcaatcgaacagctacagctttagggccagagaa 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



Db 1180 agtagctgaagaggtctctggcagcaattagaaactgtagcttgcattcttgaggagcaaaagaa 1239  
OY 639 agaacttaagaaggtatcacagaactctcaagaatgcaagatttggcataaaaggac 698  
Db 1240 agaacttgaagaaggtacacaaaatttagaagaagctaaagaatttggataaagaagc 1299  
OY 699 tatagctcaaaaggtctctcttggtgtgtactctcttgaaatggaactagagct 758  
Db 1300 tatcacagcaaatattctataggtgcgtctctcctgtacatctatgactatgactct 1359  
OY 759 tgccttttgtagtggaaacctcttgattcttaattggaagaactggaatatacatcgagac 818  
Db 1360 ggcctctgtagtaggaccacttgctctctcagggg-----atatctattgggaa 1413  
OY 819 tgtctctgcttcttctttagtgaatccatagcaggtattgcaattggagcagcagctcc 878  
Db 1414 agtactcctgtatctcttcttgtagtlaattggtggtctttagtgttgagcagcgctccc 1473  
OY 879 tcactttgaaacctctcgaaatagagccgagaggtctgcttcaatttccaggtattga 938  
Db 1474 aagcttgaagaacttgcgaatgcgaagaagcagctatgaaacttcaagaatattga 1533  
OY 939 taagaaaaccaggtataga taaacttccacagcttgatataaactgataccatagaag 998  
Db 1534 taataagccaaggtattgacagctatctgaagaatggtggcacaacccaagataattaaagg 1593  
OY 999 aactgtgaatttaaaaatgttcttcaattatccatcaagaacccatctatcaagattct 1058  
Db 1594 aaatttggaaatcggaaatgttccacttcaagttacccactcggaaaaagatlaagatctt 1653  
OY 1059 gaaaggtctgaatctcgaatlaagcttggagagacagtcgcttggtgcgtgtcaatgg 1118  
Db 1654 gaagggcctgaaccttgaaagtgcgaagtcggcgaagcgtggtggcgttggaaaaagtcgg 1713  
OY 1119 cagttggagaagtagcgtgtagtccagctcttcgagaagttatagatccgagtagtcgtt 1178  
Db 1714 ctgtggagaagcgaacaacagtcagctgtagaggtctatagccccacaaaggaggt 1773  
OY 1179 tatcatgtgtagatgaatagatagacagctttaaatgtcgagcatctcgagaccatat 1238  
Db 1774 ggtcaggtctgtagacacggaatattagacataaattglaaggttcttcaagggaaatcat 1833  
OY 1239 tggaggtgttagtcaagagcgtgttcttgcggagaccctccagttaacataatagta 1298  
Db 1834 tgggtgtgtgtcgcaggaactcgtgttcttgcaccaacagtagctgaanaacatcgcta 1893  
OY 1299 tggacgagaatgtagtgcagtagaagaatggaagacgaagcgaaggaacatgcgta 1358  
Db 1894 tggcgtgtaaaatgtaaccatgtagatgtagattcgagaagcgttcaaggagccaatgccta 1953  
OY 1359 tgaatttatcatgtagtttctctaataaatttaatacatctgtaggaggaanaaggagctca 1418  
Db 1954 tgaacttatcatgaaacgcgctcataaattgcacacccctgtgtggaagagaagggccca 2013  
OY 1419 aatgagtggaagcgagaaacagagatcgaaattcgtgcgttcaattagctgaaccccaa 1478  
Db 2014 gttgaagtgtgtggcagaagcagagatcgccaattgcaagtcgctcggtctgcaaccccaa 2073  
OY 1479 gattctgaatttagatgaggtcagctgcgtccttgattcagaanaagcaagtcagctgtca 1538  
Db 2074 gatccctcgtcgtgtagtgagccacgltcagcccttgacacagaagaagcgaagtggttca 2133  
OY 1539 agctgcacatggaagaagcgaaacaaagtccggaactacaatcgtgtgtagcacaccgacttc 1598  
Db 2134 ggtggtccttgataaagcgcaaaaaggtcgaaccacacttgtagtagctacgttctgc 2193  
OY 1599 tactattgaagtgagattgtagtttgacccctaaagaatggaatgagcggcgcgcaagaag 1658  
Db 2194 taacgttgttaatgctgcgtcatcgcgtgttgcgtagatgaggtcatctgtggaaaaagg 2253  
OY 1659 agcacatgctgaactaatgycanaaacgaaggtctatatattatcaactgtgtatgctcaaga 1718  
Db 2254 aaatcatgtagtaactcatgaaagaagaagcatcttactcaaacctgtgcacaatgcagac 2313  
OY 1719 tattaaaaagctgtagaacagatgtagtcaatgcaattcta----- 1762  
Db 2314 agcaggaatgtgaagtttgaatttagaaaatgcaagctgtagtgaatccaaaagtgaatgtagtc 2373  
OY 1763 -ctgaagaagaagccaactcctcctgcagctctgtgaaagagcatca----- 1810  
Db 2374 ctggaaatgcttccaatgattcaagatccagctcttaataagaanaagtaacctcgtag 2433  
OY 1811 -----agtcagacttcatgacaaggtctgaggaaatccaccatcaagaagat 1859  
Db 2434 gagtgcgtgtagtacaagcccaagacagaagaagcttagtaccagaagcgctcgtgtaga 2493  
OY 1860 aagcttccctgaagctcctctatataaattttaaagttaaaccaagcctgtagtcctt 1919  
Db 2494 aagtaactccaggttcccttcttgtaggaatttgaagaaatttaactgaaatgtagctta 2553  
OY 1920 tgtgtctcggagacattggtcctgtcttaagtgaactgttcaaccagtaattccat 1979  
Db 2554 ttgtgtgtgtgtatcttgtgcatataaattggaaggtcgtgcaaccagcatgtgcaat 2613  
OY 1980 catcttgcanaaattalaacacatgtt---ggaataatgatalaaaccacatlaaagca 2036  
Db 2614 aatatctcaagaattataggggttttacaagaagttagtgatgctcctgtaaaacaaagca 2673  
OY 2037 tgaatgcagaatattatccatgatatctgcatacttgggtgtatctgtctgtcaatga 2096  
Db 2674 gaatgtaactgttttactactatgttctcagcccttggaaattatcttcttatcaact 2733  
OY 2097 ttctcagagagattattttagcagagcgagggaaattttaaagcagtagaattagaaca 2156  
Db 2734 ttccctcaaggtttccaacttggcaaacgtagagagatccctcacaagcggtccgata 2793  
OY 2157 ctggcttccaagcaltgatalacagataltgcgtgtgttga tggaaagyaaaagaa 2216  
Db 2794 catggttttcgatcatcagctcagacagagatgtggtgttggatggcccttaaaacac 2853  
OY 2217 cacaagaagcttgacaacaatatatagccaatagatatagcacaaattcaaggagcaagg 2276  
Db 2854 cacttgaagcatgtataccaggtctgcgaatgtctgtcgaagttaaagggtccatag 2913  
OY 2277 ttccaggaattggtcgttcaacaacaatgcaactacaactgagacttccagttatcatctc 2336  
Db 2914 ttccaggtctgtcaatataccagaatatagcaaatcttgggaacagaatatatatatc 2973  
OY 2337 ctlatatagatggagatgacatctccgtatctgtagtaltgtctcagtaactgtgcgt 2396  
Db 2974 ctcatcatggttgcacaactaacactgttactcttagcaattgtacccatcatgcaat 3033  
OY 2397 gacagaatgattgaaacccgcagcatgactgattgccaacaagaatgaagaagact 2456  
Db 3034 agcaggagtlgttgaagaatgttctgcgaagaacgaactgaaatgaagaagact 3093  
OY 2457 taagcatcgtggaagatagcaactgaaagccttggagataatagctactatagttcact 2516  
Db 3094 agaaggtcgtgggaagatcgcctacgtgaagcaatagaanaacttcgaaacggtgttctt 3153  
OY 2517 aacaaaggaaaaagcctcogagcaaatgtatgaagagatgcttcagactcaacagaaa 2576  
Db 3154 gactcagagagcagaagtttgaacatagtagtctcagaggttgcaggtacacatacagaaa 3213  
OY 2577 taacctgaaagaagacagatattatggaaagcgtgtatgattcaacatgctcttatata 2636  
Db 3214 ctcttgggaagacacacatcttggaaattacatttctccaccacggacaatgtagta 3273  
OY 2637 ttctgcctatgacagaggttgcgaattggagaccttaattcaagctggaagcagatgac 2696  
Db 3274 ttcttccatgctgtagtattccggttggagaccttaacttggtagcatataaactcaatgag 3333  
OY 2697 cccaagagcagtgtcatagtttctactgcaattgacatagtagagctatgccaatgcgaa 2756  
Db 3334 cttagagagtgctcgttagtatattcagcgtgttcttggtagcacaagccgtggggca 3393

QY 2757 aacgctcgttttggctcctgaatatccaaagccaatcgggctgcgcatctgttctgc 2816  
 Db 3394 agtcaagttatcttcctccgcatatgcgaagcaaatatatacagccacatcat 3453  
 QY 2817 ctgtgtggaagaagaaccctaataatagacgcgcagtcgaagaaggaagacacac 2876  
 Db 3454 gacatctgaaataaaccccttctgattgacagctacgaacgaagcgcttaatgcgaacac 3513  
 QY 2877 atgtgaagggaatttgaagtttcgaagaagtccttcttcttccatctgcgcgaatgt 2936  
 Db 3514 attggaaggaataatgatacatttggtagaagttgattcctaactatccaccgcgaacat 3573  
 QY 2937 ttcaatcctccgtgtgcttaccctcaatgtagcggaagaagacagtagcatlttggtg 2996  
 Db 3574 cccagtgcttcaggagactgagcttggaagtggaagggcagcagctgtgtgtgtg 3633  
 QY 2997 gagcagcggtgtgtggaagaagcactctctgtcgaactctcgaagacttatacccgct 3056  
 Db 3634 cagcagtggtctgtggaagagcagtggtccagctcctcgtgagcggttctacgaacctt 3693  
 QY 3057 gcaaggaacaagtgtcttltgattgtgtgtagtgcaaaagaaattgaatgtacagtggtccg 3116  
 Db 3694 ggcaggggaagatgtctgtgtgtagcgaagaataaagcgactgaatgttcagtggtccg 3753  
 QY 3117 ttcccaaatagcaatcgttcccaagaagcctgtgtcttcaactgcgcattgtgtgaa 3176  
 Db 3754 agcaacacctgtggtcgtgtgcacagagccacatcctgttgcactgcagcatgtgtgaa 3813  
 QY 3177 catcgccctatgtgtggaacaacgcgtgtgtgtgccttagagtagatcaagaagcgcga 3216  
 Db 3814 catgtccataggaagacaacgcgcgtgtgtgtcacaagaagagatcgtgtgggagcaaaa 3873  
 QY 3237 tgcagcaaatatcatcttcttcttgaaggtctcctgagaagaatacacacaagaatttg 3296  
 Db 3874 gggagccaacatacatatccttcatcagtcacgtcctaataatagactaagaatgag 3933  
 QY 3297 actgaaagagcagacagcttctcgtgcgcgaagaacaagaactatcttgaaggtctc 3356  
 Db 3934 agacaagaagactcagctctctgtgtgcgaagaacaagcatltgcatagtctgtgcct 3993  
 QY 3357 tctccaaaagcccaaaatttatgttggatggaagcagctcagccctgcgaataagcag 3416  
 Db 3994 tgttagacagcctcatatatttgccttctgtagaagcagtcacgtccttgatcagaaga 4053  
 QY 3417 tgaagaagtggttgcagatgcgccttgaataagccaagacggaaggaagacagctagtg 3476  
 Db 4054 tgaagaagtggttgcagatgcgccttgaataagccaagacggaagacggaagacagctagtg 4113  
 QY 3477 caatcagagctctctcgaatcagaagcagatlttgatagtttctgcacaatgaa 3536  
 Db 4114 tgcacacgcctgttccacacacaggaatgcagacttaatagttgttctcagaatgag 4173  
 QY 3537 gataaagaagaagaagaaactcacaagcgtcctcgaagaatcgaagacataatttaagt 3596  
 Db 4174 agtcaagaagacatgagcagcagcagcgtgtgtgcacagaagaagcatctatttcaat 4233  
 QY 3597 agtgaatgcacagtcag 3613  
 Db 4234 ggtcagtgctcaggtcgt 4250  
 RESULT 9  
 AAT13394  
 ID AAT13394 standard; DNA; 6505 BP.  
 XX  
 AC AAT13394;  
 XX  
 DT 24-JUN-1996 (first entry)  
 XX  
 DE Hybrid vector pSF-MDR.  
 XX  
 KW Hybrid; vector; gene transfer; gene therapy; haematopoietic stem cell; retroviral; murine embryonic stem cell virus; MESV;

KW Moloney murine sarcoma virus; (MOMuSV);  
 KW Friend murine leukaemia virus; F-MuLV; ds.  
 XX  
 OS Synthetic.  
 XX  
 PN DE19503952-A1.  
 XX  
 PD 14-MAR-1996.  
 XX  
 PF 07-FEB-1995; 95DE-1003952.  
 XX  
 PR 08-SEP-1994; 94DE-1431973.  
 XX  
 PA (BOE ) BOEHRINGER MANNHEIM GMBH.  
 PI Baum C, Osterlag W, Stocking-harders C, Stockingharders C;  
 DR WPI; 1996-152306/16.  
 XX  
 PT Hybrid retroviral vectors - for gene transfer into haematopoietic  
 PS stem cells  
 XX  
 PS Disclosure; Page 25-29; 42pp; German.  
 CC  
 CC New hybrid vectors comprise (1) a leader region including the U5  
 CC region and TRN primer binding site of murine embryonic stem cell  
 CC virus (MESV) or Moloney murine sarcoma virus (MOMuSV), and (2) a 3'-  
 CC LTR including the U3 and R regions of a Friend murine leukaemia  
 CC virus (F-MuLV). The vectors are useful for ex-vivo or in-vivo gene  
 CC therapy. High levels of gene transfer can be achieved in  
 CC haematopoietic stem cells and their myeloid (non-lymphatic) progeny.  
 CC pSP1, pSP2, pSP3 and pM1 (sequences given in AAT13390-T13393) are  
 CC examples of such vectors.  
 CC Vector pSF-MDR (sequence given in AAT13394) is based on the  
 CC MESV vector R224.  
 XX  
 SO Sequence 6505 BP; 1719 A; 1506 C; 1606 G; 1674 T; 0 other;  
 Query Match 36.38; Score 1316.2; DB 17; Length 6505;  
 Best Local Similarity 62.3%; Pred. No. 0;  
 Matches 2178; Conservative 0; Mismatches 1268; Indels 51; Gaps 5;  
 QY 162 caggttaccctgtatattatgttgaataagtggtgcgccttatttggtaacata 221  
 Db 2152 caggtatgccattatcatcagtgtaattggtcgggggtgctgtgtcgtcatatca 2211  
 QY 222 gatttcctgtgattataactgcagcagcagacccaagaggaattcgaaacagtttt 281  
 Db 2212 gtttcatlttgggtgctgcgcagctggaagcaataacaaattagaacagttttt 2271  
 QY 282 tcaatcagtttggcagacagacatcgcgtgtgtgatagcgtgtgacatcggtaactaa 341  
 Db 2272 tcatgataatgacagagagatagcgtgtgtgtatgtgcagatgttggggagactaa 2331  
 QY 342 cactcgcgtgcacacatgac-----aaatacagatgattgtgaataagatgtcct 398  
 Db 2332 caccgcgacttacaatgattgtctcttaagataagaaattatgtgtgacaaattggaat 2391  
 QY 399 gttgttcaaaacatgtctacttcttcgatcgtgcgcgtgcaattgttgggaagggctg 458  
 Db 2392 gttcttcaagcaatggaacatttctcagtggtttatagaatttaacagtggtgt 2451  
 QY 459 gaactcacctatgtactatccacagctctcctcttaataatgagcttaacggcagcatg 518  
 Db 2452 gaagtaacctgtgtgatttggccatcagctgttcttgagctgtaagcgtcgtctg 2511  
 QY 519 ttctagagatgtcatcctatgtgcaggaagaaatgaatgtcctatttccaaagcggtggc 578  
 Db 2512 ggcacaagaatactatcttattcagtaaaagaaacttlaacgtatgcaaaagcggtgagc 2571  
 QY 579 tgtggcagaagaagctgtgatacattccgaacagtcatacgtttagggcccaagaa 638

Db 2572 agtagctgaagaagctcttgccagcaattagaaactgtagttgatttgccattgtagagacaaagaa 2631  
QY 639 agaacttcaaaaggtatacacgaatctcaagaatgataagaattttgcatataaaaggac 698  
Db 2632 agaacttgaagaagtacacaaaataatttagaagaactaaagaattggaatataaagaagac 2691  
QY 699 tatagtccaaaggtcgtctggtcgtgtagtactctttaagattggaattggaacttggact 758  
Db 2692 tattcacgccaattattctatagtagtgcgtcttctcgtgactctatgactctatgactt 2751  
QY 759 tgccttttgtagtagaacctcctctgactcttcaatgtagaaccctgtagatatacatcgggac 818  
Db 2752 ggcctcttgtagtagaaccaacttggctctcctccagggg-----atatctatttggaca 2805  
QY 819 tgtctctgctgttcttcttagttagtataccatagcaattatgcatltggagcagcagctcc 878  
Db 2806 agtactcctgtatcttcttctctgtatlaatttgggctttagtltgtgacagagcactcc 2865  
QY 879 tcaacttgaaacctcgcgaatagcccgagagctgaccttcaatttcttccaggtatttga 938  
Db 2866 aagcatgtgaagcatttgcgaatagcaagtagagagcactatgaaatcttcaagatatgtga 2925  
QY 939 taagaacaaccgtagatagaacttctccacagctlgatataaaccctgatatccatagaag 998  
Db 2926 taataagccaagtattgtagcaggtattcgaaagtgtaggcacaacaacagatatataaggg 2985  
QY 999 aactgtggaatttaaaaatgttcttcaatataccatcaagaacatatataagattct 1058  
Db 2986 aaatttgaatcagaataatgttcaacttcagttaccatctcgaaaagaagtataagattct 3045  
QY 1059 gaaagctcgaatcagaatlaagctcgaagtagagacagctgcgtcttgcgtctcaattg 1118  
Db 3046 gaagagcctgaaccggaaggtgtagagtcggcgagacggttgcgtctgttggaaacaagtgg 3105  
QY 1119 cagtggagaagtagtcgtagtccagctcttcgagaggtatatatgatccggaatgagctt 1178  
Db 3106 ctgtggagagagcacaacagctcagctgtagcagaggtctctatgagccccaagagggt 3165  
QY 1179 tatectgtgtgtgtagaattgatactagaagctttaaattgtgcgcaattatcgagacatt 1238  
Db 3166 ggtcaggtgtgtagacagatataagaccaaataatgaaaggttcttctacgtaggaatct 3225  
QY 1239 tggaggtttagtcagagcgctgttcttgcgggacacacagtaaacatatacaagta 1298  
Db 3226 tgggtgtgtgtagtcagagacccgtatgttctgcacccaagataagcttgaaataactcgca 3285  
QY 1299 tggagagatagtagtcgtagatgtagaagatgtagagagcagcagaaggaagcaatgcyta 1358  
Db 3286 tggcgttgaaaaatgtccacatlgatgagatttgaaaagctgtccaaaggaagccaattgcta 3345  
QY 1359 tgaatttatcatgtaggttctctataaatttaatacatgtgtgtaggggaaaaagagctca 1418  
Db 3346 tgaacttcatgaaactgtcctcatataaatttgcacccctggttggagagagggccca 3405  
QY 1419 aatggttgagggcgagaaacagagatgcgaattgctcgtctgacttaagtctgaaacccca 1478  
Db 3406 gttgaggttgaggcgagagcagagatgcgcactgvcacgtgccttggttgcgaacccca 3465  
QY 1479 gattctgattttagatgtaggtactgtctgccttgattatagaagaacgaatlcagctgttca 1538  
Db 3466 gatctctcgtcgtatgtagggccaagctgcagcttgtagacaagaagaagcgaagcagtggttca 3525  
QY 1539 agctgcacttgagaagcgagcagaaggtcgcgactacaatcgtgtagcacacgacttcc 1598  
Db 3526 ggtgctcctcgtgaataagcgagaaaaggtcgcgacacgaattgtagctcatcgttctgc 3585  
QY 1599 tactcttgaagaattgagattgattgtgacccctaaaggtggaagtgtgtagggagaag 1658  
Db 3586 taagcttgtagtgcgtgcagctcagctcgtgttgcagtgtgtagagtcattgttggagaag 3645  
QY 1659 agcaaatgtcgaactatgccaacagaggtctatataattcaactgtgtatgcagaaga 1718  
Db 3646 aatcatgtatgacatcattgaagaagagaagcatttaacttcaaacattgtcacaaatgcagac 3705

---

QY 1719 tattaaaaagctgtatgaacagatgtaggttcaattgacatatctca----- 1762  
Db 3706 agcaggaatattgaattggaattagaaatgcagctgatgtagatcccaaaagtgaattgatgc 3765  
QY 1763 -ctgaagaagaagaccaactcacttccctcgcactctgtgaagagcatca----- 1810  
Db 3766 ctgtgaaatgtcttcaaatgattccaagatcccgctcttaataaagaaaaagataacccgtag 3825  
QY 1811 -----agtcagacttcatgacaagctgtaggaatccaccaccaatcaagaagat 1859  
Db 3826 gagtgcgtgtagatcacaaagcccaagacagaagcttagtaccaaaagagcctgtgatga 3885  
QY 1860 aagcttccctgtagtctctctatttaaatttaagttaaacaagcctgtagtctt 1919  
Db 3886 aagttacctccagttcccttlttgtagagatttagaagcttaaaattcaactgaaatgcttca 3945  
QY 1920 tgtgttctcgggagacatttgcctcgtctcaatggaactgttcatccagtaattccat 1979  
Db 3946 ttgttgttgtgtcatlttgtgccattataatgtaggagcctgcgaacccagcatttgcgaat 4005  
QY 1980 catcttgcgaataatataaccaatgtt--ggaataatgataaaaccacataagca 2036  
Db 4006 aaattttcaaaagattataggggtttttacaagaatttgaatgtatccctgaaacaaagca 4065  
QY 2037 tgaatcagaatattatccatgatatttgcatttttgggtgtatttggcttgtcatgtta 2096  
Db 4066 gaattgaactgttcttcaactatgttcttagcccttggaaattatcttcttattacatt 4125  
QY 2097 ttctacgtcagagattattttacgcgagacagaggaatttaacatgatatagata 2156  
Db 4126 ttctctcagagtttccacttgcgaagctgtagagatccctcaccaagcgtctcgata 4185  
QY 2157 ctggccttcaaaagcatgtataltacagatatgtccgtgtttagatgaaaggaacaag 2216  
Db 4186 catggtttccgcatcatcagatcagacaagatgttggttgtttagatgacccaataaacac 4245  
QY 2217 ccagagagcttgcacaacatatattgcccataatagacaaatccaagggccaaggg 2276  
Db 4246 cactgtagcatgtgataccaagctgcgcaatgatagtctgcctcaagttaaagggcatalagg 4305  
QY 2277 ttccagagattggtcgttcaacaaaatgcacaactcaacaggaacttcaagttatcatc 2336  
Db 4306 ttccaggtctgtgtaattaccagaatatagcaaatcttggagacagatatattatattc 4365  
QY 2337 cttaataatgtagtggagatgacattccgtatcttcbagttatgtctccagttgcgt 2396  
Db 4366 cttaactatgttggcaactaacatgtaacttactcttagcaattgttaaccatattgcaat 4425  
QY 2397 gacaggaatgtatgaaacccgcagcaatgactgtgatttgcacaacaaagataagcaagaact 2456  
Db 4426 agcagagattgttgaataatgaaaatgttctgtagacaagcactgaaagataaagaagaact 4485  
QY 2457 taagcatgtcgtgaaagatgacacaggaacttggagaatatagctactatagtgctact 2516  
Db 4486 agaaaggtcgttggaaagatcgtcactcgaagcaatagaaaacttccgaacccgttgttctt 4545  
QY 2517 aacaagggaaaaagccttcgagcaaatgtatgaagaaatgcttccagactcaacacagaaa 2576  
Db 4546 gactcagagagcagaagtttgaacatatgtatgtctcagagtttgcaggtatgcatacagaaa 4605  
QY 2577 taactcgaagaagaacagatatgtggaagctgttatgtatccagccaatgctttatata 2636  
Db 4606 cctctttagaagaagcacaactcttggaaattacatttcttccacccaaggaatgtagta 4665  
QY 2637 ttctgcctatgacagaggttgcgaatttggagccttaattcaagcgtggaagtagac 2696  
Db 4666 ttcttccatgtcgtgatttccggtttagagcctactcgttggacacataaactcatag 4725  
QY 2697 ccagagagcatgttcatatttacttactgcaattgtcataatgtagagctatgtagccatcgaaa 2756  
Db 4726 cttttagagatgtctcgttagtatttccagctgttgccttggltgcattgscgttggggaca 4785

[illegible]

OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CD	424..4267
FT	/tag= a
XX	
PN	W09324613-A.
XX	
PD	09-DEC-1993.
PF	
XX	
PR	14-MAY-1993; 93WO-USO4707.
XX	
PA	22-MAY-1992; 92US-0887712.
XX	
PI	(GENE-) GENETIC THERAPY INC. (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX	
DR	Mcdonagh KT, Nienhuis A, Tolstoshev P; P-PSDB; AAR44297.
XX	
PT	WPI: 1993-405805/50. P-PSDB; AAR44297.
XX	
PT	DNA or RNA sequence for human multi-drug resistant gene MDRI - encodes P-glycoprotein with altered splice site, used in cancer therapy
XX	
FS	Example: Fig 4; 64pp: English.
XX	
CC	PMR2000 contains an mdrl cDNA sequence (AA052126) described in PCT application no. WO87/0943m wherein the first 282 bp of the 5' UTR and the last 23 bp of the 3' UTR of the cDNA sequence have been removed. The codon YTN is unclear in the spec., and encodes the AA Leu.
CC	
XX	
SQ	Sequence 4669 BP; 1394 A; 893 C; 1126 G; 1254 T; 2 other;
Query Match	36.3%; Score 1314.8; DB 14; Length 4669;
Best Local Similarity	62.2%; Pred. No. 0;
Matches 1176; Conservative	1; Mismatches 1269; Indels 51; Gaps 5;
QY	162 caggttgaccctgatattatcgttgaataagttgctgccttgatttttgttacataca 221
Db	760 cagatcagtcattattacagtgaattgtgctggggtcgtgctccgtctacattca 819
QY	222 gatctccctgtgattatatcactgcgcgcagaacagcaaggatttcgaaacagtttt 281
Db	820 ggtttcatlittgttcctcgccgtgcaagaaacaatacacaaaattagaacacgtlitt 879
QY	282 tcatcagatttggcgcacagacatcgccgcgtgttgatcgttgacatcgctgaacttaa 341
Db	880 tcatgtcataatgvcagcagagagatagcctggtttgatgtgcacgattgttggcagcttaa 939
QY	342 cacctgcatacagacatctgac--aaatcacgtgatgtgatcttgagataagattgctct 398
Db	940 caaccgacttacagatgattgctctcctaagttaatgaccttatgttgcacaaatttgaaat 999
QY	399 gtgtttccaacaacatgctactcttttcgattgctcctgcgcagttgtgttaaggcgc 458
Db	1000 gtcttttcagtcataatgvcacacatlittcacctggtttatagataggaattcacacgttggtgc 1059
QY	459 gaactcacccctcagtgbaectctacacgcgtcctctctatataatgcttcagcggcagacatg 518
Db	1060 gaagctaaccccttgatattttggccatcagctcctctgttcttggaatcgtcagcgtcgtctgc 1119
QY	519 ttctcgaatgctcatalcctcatcttgaccgaataaggaaatgaatgaatctccaaagcttgggc 578
Db	1120 ggcaagaatcattatctctcttaccatgataaagaacaccttagcgtatgcacaaagctggagc 1179
QY	579 tgttgccagaagaagctcttgttcatcaatccgaacagatcagaagctttaggtcccagagaa 638
Db	1180 agtactcgtgaagggtcttcttgcagcaattgaacttgatgattcttgtagggacaaagaa 1239

[illegible][illegible]

```

Db 3394 agtcagttcattgtcctcgtactatgccaagccaataatcagcagccacatcatcat 3453
Qy 2817 ctgttggaaaagaacaaatatagacagccgagtcacaagaaggaaaagccagacac 2876
Db 3454 gatcatgaaaaccccccttgaattgacagactacagacaggaagcgaatgacgaacac 3513
Qy 2877 atggaaggaattagatgattcgaagaatctcttcttctacatcatalgtcgccagatgt 2936
Db 3514 attgaaaggaaatgcatcatttgtagaagtgtatctacatccaccgacccgagacat 3573
Qy 2937 ttcatccctccgtgtgtatctccctcagatattgagcagagaaagacagatgattgtgg 2996
Db 3574 cccagttctcagagactgagcctgtgaggtgaagaagggccagacgctgtcctgtgtgg 3633
Qy 2997 gaggcagggcgtgtggaaaagcactctgtctcaactcttcagagacttcttgcccgct 3056
Db 3634 cagcagtggtcgtgtggagacacagtggtccagctccttgagagcgtgtctcagacccct 3693
Qy 3057 gcaaggaacaagtgtgttggtagtgatgacaaaagaatgaaatgacagtggtcctcg 3116
Db 3694 ggcagggaaaagtgtcgtgtgtagcaagaataaagcagtaattgttcagtggtcctcg 3753
Qy 3117 ttcccaaatagcaatcggtctcacaagacgtgtcctctcaactgacagcattgtcgaa 3176
Db 3754 agcacaccctggcactgctgtcccaaggccacatcctgtttgactgacagcattgtcgaa 3813
Qy 3177 catgacctatgtgacaaacagccgtgtgtgacattagatagagatcaaaagccgcaaa 3236
Db 3814 catgtccctatgagacacacacccgggtgtgtcacaagaagacgtgagggcgcgcaaa 3873
Qy 3237 tgcagcaaatatccatcttatttgaaggtctccctgagaataacacacacacagttgg 3296
Db 3874 ggaagccacaatacattgctctcatcgagtcactgctcaataataatagcactaaagtgg 3933
Qy 3297 actgaaggagacacagcttcttgccgcccagaaacaaagactagctatttgaaggctct 3356
Db 3934 agacaaaagaaactcagctctctgtgtgccaagaacacagcattgacagctgtgcct 3993
Qy 3357 tctccaaaacccaaaatttatgttgaatgagagccacttcagccctcgagtaagacag 3416
Db 3994 tgttagaagacctcatatttggcttcttgatgaaagcagcagctcagctcgatacaagaag 4053
Qy 3417 tgaagaagttggttcagcagctcgtgataaagccagacagcgggaagacatgcttagtgt 3476
Db 4054 tgaaaagttgtccaaagagcccttgacaaagccagagaagccgacactgtcatgtgat 4113
Qy 3477 cactcagcagctctctgtaattcagaacgcagattgtatgtgtctgcacaatggaaa 3536
Db 4114 tgcaccgcctgtccaccatccagaaatgcagacttaattgtgtgttcacagaatggcag 4173
Qy 3537 gataaaggaacaagaactcatcaagagctcctgagaataatcgagacatatatttaagt 3596
Db 4174 agtcaagagatgtgcaagcatcagcagctgtctgtgacacagaagaagcatttttcaat 4233
Qy 3597 agtgaatgcacagtcag 3613
Db 4234 ggtcagttgtccaagctg 4250

```

```

KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KM cytokine; wild-type; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3843
FT     /*tag= a
FT     /product= "Human wild-type MDR-1 protein"
FT     replace (553..555, GTT)
FT mutation
FT     /*tag= b
FT     /note= "cDNA sequence of G185V human mutant MDR-1 given
XX     in AA249333"
XX
PN W09961589-A2.
XX
PD 02-DEC-1999.
XX
PE 27-MAY-1999; 99WO-US11825.
XX
PR 28-MAY-1998; 98US-0086988.
XX
PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Sorrentino B, Bunting K;
XX
DR WPI: 2000-072615/06.
DR P-PSDB: AAY58186.
XX
PT Ex vivo expansion of hematopoietic stem cells transduced with a
PT sequence encoding human multidrug resistance-1, used for bone marrow
PT transplantation -
XX
PS Claim 10: Page 68-70; 113pp; English.
XX
CC This sequence represents cDNA encoding human wild-type
CC multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC efflux pump, responsible for the export of drugs from cells,
CC particularly cancer cells. Wild-type MDR-1 shows increased
CC resistance to etoposide and decreased resistance to vinca
CC alkaloids compared with a mutant form (AAY58187) where the Gly at
CC position 185 is replaced by Val. The invention relates to transducing
CC hematopoietic stem cells with nucleic acid encoding an MDR protein
CC and culturing the modified cells. The modified hematopoietic stem
CC cells are useful in bone marrow transplantation (to reconstitute
CC hematopoietic systems in patients who have undergone chemotherapy or
CC radiation therapy) and in ex vivo gene therapy of genetic defects in
CC cells derived from hematopoietic stem cells, e.g., thalassemia,
CC Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC cells can also be used to identify factors involved in regulating
CC proliferation and differentiation in hematopoietic stem cells.
CC Hematopoietic stem cells that express MDR-1 will be protected against
CC chemotherapeutic agents, so can be engrafted while the patient is
CC undergoing chemotherapy. Expansion of (rare) hematopoietic stem cells
CC provides sufficient cells to permit standard biochemical analysis.
CC Overexpression of MDR-1 allows cytokine-driven expansion of
CC hematopoietic stem cells by at least 10-fold compared with a maximum
CC of 4-fold in known procedures.
XX
SQ Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

```

```

Query Match 36.3%; Score 1314.6; DB 21; Length 3860;
Best Local Similarity 62.3%; Pred. No. 0;
Matches 2177; Conservative 0; Mismatches 1269; Indels 51; Gaps 5;

Qy 162 caggttaccctgtatattgttgaatagggtgtcctcctgattttgtgtacatca 221
Db 336 cagatagctattattacagtggaatgtgtcgtgggtgtcgtgtgtcgtacatca 395

Qy 222 gattcctgtgtgattataactgcagcagcagacagcagaagagattgaaaacatttt 281
Db 396 gtttcatttgtgtcctgtgcagctcgagaacaaataacaaattgaagaacagtttt 455

```





Db	2610	agcgaggggtctgttgcgaatgtgaaatgtcttgcctgcggacaagcaccggaagataagaaagaaact	2669
OY	2457	taagcaatctctgtagaagaatagcaaatctgaaagcctcttgtagaataatagctactatagcttcat	2516
Db	2670	agaaggtcgtcggtagaagctcgctacttgaaagcaatagaagaaacttcgaacctgtgttctctt	2729
OY	2517	aacaaaggagaaaagcctctcgagcaaatgtatagaagagatgcttcagactcaacacagaaa	2576
Db	2730	gacctaaaggaaacaaagttttaaataatgtatgtctcagagatttcgaggtttaccatacgaagaa	2789
OY	2577	tacctcgaaagaagcacaagatltatgtgaagcgttatgtgatttcagccatgcttata	2636
Db	2790	ctctcttgaggaaagcaacaacactcttggaatlaacatttccctcacccagcgcaatgata	2849
OY	2637	ttttgcctatagcaagcaaggttttcgaattcggagcctatttaattcaagcctgtagcaatgac	2696
Db	2850	ttttcctatgctgtagatgttcctcggttgtagagcctactctggtgagcaataaactatagag	2909
OY	2657	cccaagagcgatgcttcaatagtttttaactgcgaatgtcatagtgagctatgcccacgaaa	2756
Db	2910	ctttgagagatgtctgtttagtattttcactcagctgtttgctttgtgcacatgcccgttgggca	2969
OY	2757	aagcgtcgttttgcctcttcataatcccaagcacaatccggtggtctgcgatactgttgc	2816
Db	2970	agtcagttcatctgtcctcgcgaactatgcgaagcacaataatcacagaccccaatcaatcat	3029
OY	2817	ctcttgtagaaagaaaacaataataagaacgcgcacatcaagaagaggaagaaaagccagcaac	2876
Db	3030	gactcatgtgaaaaaaccccttgttgtagacagctacagcacaagcagaagcctcaatgtccgaacac	3089
OY	2877	atgtgaagggaaatttagagtttcgagaagctctcttctctatccatgtgcgccagatgt	2936
Db	3090	attgaaaggaaatgtccacatcttggtagaagttgtatgtacatactccacccgcagacat	3149
OY	2937	tttcatcctcgcgtgtactccctcagctatgttagcgaggaagaaagacagatgttggtg	2996
Db	3150	cccaagctctcaagagactcgagccctgcgagagtgaaaggaagccagacgcgtgcctctgggg	3209
OY	2997	gagcagcgcgcgtctgtagaagaagcactctcgttcacactctgcagagaaactttatgaaacctg	3056
Db	3210	cagcagatggtcgttgtagaagcacaagtggtccagcctccctcgagcggttctaagaaacctt	3269
OY	3057	gcaaaggacaagctgcttcttgaatggtctgtagatgcgcaaaagaatgtaactgtacagttgctcg	3116
Db	3270	ggcagctgaagaagtgctgtcttgaatgcaaaagcaataaagcgacgtgaattgttcagttgctcg	3329
OY	3117	ttcccaatatgcaatcglttctctaaagaagcgtgtgccttcaactctgacagatgttgcgaa	3176
Db	3330	agcacacacttgagcatctgtgtcccaagagaccacactcgttttgactgcgaagatgtcgtga	3389
OY	3177	catgcgtctatgtgtacaacaacgcgcgtctgtggtgcattatagaatgataaagaagccgcgaaa	3236
Db	3390	catgtcctatgtagagacaacagccgggtgtgtctacaagagaaagatcgtgtgaggtgcagaaa	3449
OY	3237	tgcaagcaaatatccatctctttaaactgaaaggtctccctcgagaataacaacaacaaagttgtg	3296
Db	3450	ggagggcaacaatacatgctcttcatctgcatgcatcgtcctaataatataagacataaagtag	3509
OY	3297	actgaaaggagacaggtttctctgtcgcgccaagaacaagaactagctatttgcaaggtcct	3356
Db	3510	agacataaggaaactatagctctctctgtgtagccaagaacaacgcatltgcataitgactcgtgcct	3569
OY	3357	ctctcaaaaaccccaaaaatttctatgtgtgtagtgagcgcaatctcagccctgcgaataagacag	3418
Db	3570	tgtttagacagcctcatatatttgccttcttgtagtagaagccacgtcagctcgtgataacgaag	3629
OY	3417	tgagaaagtggtttcaagatgacctgtgataaagcgcaagacggtgaaagacatgtgcttagtgt	3476
Db	3630	tgaaaaagttgttccaagaagcctctggaacaagcgagaaaggtccgacactgtgatatgtgt	3689
OY	3477	cactccacagctctctgcgaatttcagaagcagagatttgatagtgttctgcacaaatgaaa	3536

Accession	Sequence	Position
Dd	tgtctaccgcctgtgccaccatccagatgagacttaatagtggtgttcagaaatgacag	3749
Qy	3557 gataaagaaagaagaactcatataagaatccctcgaaataatgagacatatatttaatt	3566
Dd	3750 agtcaagaagaatgagccgcgatcaagcagctctgcgcacagaagaagcatctatttcaat	3809
Qy	3597 agtgaatgcacagttcag	3613
Dd	3810 ggtcagtggtccagctg	3826

## RESULT 12

ID AAN70752 standard; cDNA; 4669 BP.

AC AAN70752

DT 21-MAY-1991 (first entry)

Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones DE lambda-HDR10, 5 and 104.

Chemo-therapy resistant tumour cell; P-glycoprotein; ss

aa  
OS Homo sapiens.

Key	Location/Qualifiers
FH	

33

PN W08705943-A.

PD 08-OCT-1987.

26-MAR-1987

PR 01-AUG-1986; 86US-0892575.

[illegible][illegible]

XX

DR P-PSDB; AAP70452.

PT DNA for multi-drug resistance in human cells - used to detect

PT polypeptide(s) for diagnosis and therapy

PS Claim 2(a); Table 5, pp30-39; 61pp; English

CC The human multi-drug resistant KB carcinoma cell lines were used as

CC obtain cDNA clones of the *mdr1* gene (AAN70752), poly (A) and RNA was

CC that the *mdr1* gene product is likely to be a transmembrane protein

CC sites is consistent with the mdrl protein being related to the

XX

Query Match 36.38; Score 1314.6; DB 8; Length 4669;

Matches 2177; Conservative 0; Mismatches 1269; Indels 51; Gaps 5;

QY 162 caggttgacctgtattatgttgaatagtgctgtctgcctgatttttggttacataca 221

Db 760 caggtatgcctattatcagtggaattgtgtctggtgtctgtgtctgtcttaccattca 815

QY 222 gattcctgtgatatgaactgcagcagcagaccagagatcgaaaacagtttct 281



QY	2397	gacgggaatgtatgtaaaacccgcgcgaatgactgtgatttcccaacaagataagcaagaact	2456
Db	3034	 agcgggaattctgtttaaataatgaanaatgtgtctgttgcaagaacactgtgaataagaagaagact	3093
QY	2457	taagatctgtgaaagaatagaacaactgaagcttggttggaataatagctactatagatgtcatt	2516
Db	3094	 agaangtgcctggtggaagcttcgctactctgaagaataagaanaacttcgaaacctgtgttctt	3153
QY	2517	aacaaaggaaaagccttcgtagcaaatgtatgaagaatgtcttcagactcaacagaaa	2576
Db	3154	 gactaaggagagagaagatttgaacatatgtatgcagagatttgcaagttaccatacaga	3213
QY	2577	taacctggaagaagaacacaggtttatgaagctgttatgcatctacgcagctcttata	2636
Db	3214	ctcttttgaggaagaacacacalcttggaaattacatlltcccttcaccocagggaatga	3273
QY	2637	ttttgcctatgcaagcgaaggtttcgtatgttgagcctatctaactcaagctgtagatcac	2696
Db	3274	 tttttcctaagctgtagtcttcggtttgcggttctactcttggtgcaacataactcatag	3333
QY	2697	cccaagagggcatgttcaatgttttactgtcaattcatalgtgaagctatgagccatg	2766
Db	3334	 ctttgaggaatgttctgttaagtatttcaagctgtgttcttcttggccatgvcgggtgg	3393
QY	2757	aacgtctgtttgtgctctctgaatattcccaagccaactgggggctgtgcgcatctgttcg	2816
Db	3394	agtcagcttcatcttgcctctgcacatgccaagccaataataatgaagcccaatcatat	3453
QY	2817	ctgtgttgaaaagaacacaatatagacagccgcgaatcaagaagggaagaagccagac	2876
Db	3454	 gactatctaaaaaaacccctttagatctgacagctacagacaggaagcctaattgcgaac	3513
QY	2877	atgtagaagggaattagaatttgcgaagctcttctctctatcatgttgcgccagatg	2936
Db	3514	 atgtagaagggaatgtcacaatcttgtagaagtgttatccaactccacccgcgcgaagat	3573
QY	2937	tttcatcctccgtgcttactaccgtatattgagcgagagaagaagacatgtatgtgtgg	2996
Db	3574	 cccaagtgttcagggaactgagccgttgtagtgaaggaaggccagacgtctgctgtgg	3633
QY	2997	gagcagacggtcgtgtgggaagaagcaactctgttcaactcttcgcagagacttatgacccgt	3056
Db	3634	 cagcagctggtcgtgtgggaagcacaagtgtgccaagctccctggagcggttctaagacccct	3693
QY	3057	gcaaggacaagtgtgtcttatatgtgtgtagtgaagaagaattggaatgtacagttgtctcg	3116
Db	3694	 ggcagggagaagatgtgtctgttagtggcaagaataaagaagcagactgttccagtgtctcg	3733
QY	3117	ttcccaatlagcaatcgttccccaagaagccgtgtgtcttcaactcaactgcaagatgtctga	3176
Db	3754	 agcacacctgtgcatcgtgtgccagagagcccaactcgtttgactgcagatgtctgagaa	3813
QY	3177	catgcctatgtgttgaacaacgccgtgtgtgtgcattagtgtgataccaagaagccgcgaa	3236
Db	3814	 catgtccctatgtgagacaacaacgcccgtgtgtgttacaagagaagatctgtgtggcgacaa	3873
QY	3237	tgcaagcaatatccatctctttaaataaggtccctctggaataatcaacaacaagaattgg	3296
Db	3874	 ggagggccaacatatatgtcttcatctgaagtcactgcttaataatagctactaaagtag	3933
QY	3297	actgaagaagagcacagcttctctgvcggccaagaacaagaactagctatgtcaaggtct	3356
Db	3934	 agacaaagaagaactagctactctctgtgtggccaagaacaacagcatgtccatagctgtgcct	3993
QY	3357	tctccaanaaacccaatttatgtgtgtgataagagcgaacttaagccctcgataatgaag	3416
Db	3994	 tgttagacagcctcatattttgtcttgttagtagaagccaagtcagctctcgtgaacagaag	4053
QY	3417	tgagaaagtggtttcagcaagcctgtataaagccaagacgagggaagagatgctagtgt	3476
Db	4054	 tgaagaagtttctccaanaagcctgtgacaagaagccaagaagccgcactcatgtat	4113

Accession	Sequence	Position
QY 3477	caccacaggtctctcgcaatttgaagcagaattgatgaatggtgtctgcgaacatgaa	3556
Db 4114	ttgcacacgctgtgccacatccaatgaatgaactaaagaatggtgttcgaatagcag	4173
QY 3537	gataaaggacacaggaactatcaagaatctctgagaatctgagacatatatttaagt	3556
Db 4174	agtaagaagaatgacgcagatataagaagctgtctgcacagaagaagcattatlttcaat	4233
QY 3597	agtgaaatgcacagttcag	3613
Db 4234	ggtcagttgtccagctg	4250

Query Match	Best Local Similarity	Matches 2177; Conservative	Score 1314.6; Pred. No. 0; Mismatches 1269; Indels 51; Gaps 5;	DB 21; Length 8630;
162	caggttcgccccgttatctgttgcagataggggttcgtgcctgattttggtatataca	221		
1555	caggtaatgcctcatctatcaacagtggaatggggtgggtgctgctgcttacaatcca	1614		
222	gatttcctctggtgattataactcagcagcacagacagacagagattcgtaaaaacagtttt	281		
1615	gatttcattttttgttcccttcagcagctcgtgaaagaaatacaacaaaattagaataacatttt	1674		



Db 3829 agcaggagctgttgaagaatgaagaatgttcttggaacagcactgtgaagaataagaagaact 3888  
OY 2457 taagcagctgcgaagaatagaactgaagcttcttggaataataactactagtgatc 2516  
Db 3889 agaaagtgcctgggaagatcctactgaagaacataagaactccgaacccgtttcttctt 3948  
OY 2517 aacaagggaagaacccctcgcagcaaatgtatgaagaagtcttcacagctcaacagaaga 2576  
Db 3949 gactcagagcagaagatttgaacatatgtatgtcagagaatttgaggtacacacagaaga 4008  
OY 2577 taactcgaagaagaacagaatatttgaagcgtgtatgtacatccagcactcattata 2636  
Db 4009 cctcttggaagaacacacacactcttcttggaatcacatttccctccacaggaatgagta 4068  
OY 2637 ttctgtcacaagcagaggttctgacttggagcctatttaacttaacagctggaacatgac 2696  
Db 4069 ttcttctaagcttgagtggttccggttggagcctacttgggtgcacataaacccatgag 4128  
OY 2697 cccagaaggcagtgatcatttctactgcaattgcatatgagactatggccatcggaaga 2756  
Db 4129 ctttgagagatgctctgttagtaatttcagcgtgtcttcttggtgcacatggcgtgggga 4188  
OY 2757 aagcgtcgttctgctcctgaataatccaaagccaatcggggctggcactcgtttgc 2816  
Db 4189 agtcagttcatttgcctcctgactatgcacaagccaataatcacagccacatccatcat 4248  
OY 2817 ctgttggaagaagaacaaatatacagccgagtcacgaaggaagaaggaagaagcagacac 2876  
Db 4249 gatcatctgaaaaaaccccttgatctgaacagctacacaggaagcctaagtcgcgaacac 4308  
OY 2877 atgtgaagggaatttagaatttcgaagaagtccttcttctatccatgtgcgccagatgt 2936  
Db 4309 attggaagggaatttcacatttggtagagtgtatctaacatccaccccgacggagacat 4368  
OY 2937 ttcatcctccgtgtgtatccctcagatattgagcaggaagaagcagtagatcttgggg 2996  
Db 4369 cccagctcttcaggaagcagccctggaaggtgaagaagggcagcagcgtggtctgtggg 4428  
OY 2997 gaggcagcgtctgtggaagaacactctctgtcaactcttcagagagactttagcccccgt 3056  
Db 4429 cagcagtggtcgtgtggaagcagacagtggttccagctcctctggaagcgttctcgcacccct 4488  
OY 3057 gcaaggacaagtgtctgttgaagtggtgagtcgaagaagaatgtgacagtggtctcg 3116  
Db 4489 ggcaggggaagtgctgtctgtgagcaagaataaagcagtagatgtcagtgctcgcg 4548  
OY 3117 ttcccaaatagcattgttccctcaagagcgtgtctctcacaagcagcagatgtctgaa 3176  
Db 4549 agcacacctgggcacgtcgtcccgagggccacacctgttctgactgacagcattgtctgaa 4608  
OY 3177 catgcctatgtgacacaacagcgtgtgtgacatagatgagatcaaaagaagccgacaa 3236  
Db 4609 catgacctatgtgagacaacagccggtgtgtgtcacaagaagaagatcggagggacgacaa 4668  
OY 3237 tgcagcaaatatccatcttatttgaaggtctccctgagaataacacacacaaagtgg 3296  
Db 4669 ggaagccaacatatacctcattcagagtcacatcctcctaataatataagcactaaagtgg 4728  
OY 3297 actgaaggagcagacagcttctgacggccagaagaagaagaactagctatgtgaaggcct 3356  
Db 4729 agacaaggaaactcagctcctctgtgagcagaacaacagcattccatagctgtgcct 4788  
OY 3357 tctcacaagaacccaatttattctgttgaatgagggcacttcagccctcgatagacag 3416  
Db 4789 tgtcagacagcctcatatttctgttggatgagccacgctcagcctctggtacagaag 4848  
OY 3417 tgaagaaggtgttgcagatgccttgataaagcagagcaggaagcggaggaactcctagtgt 3476  
Db 4849 tgaagaaggtgttccaagaagccctggaacaagccagagaagcgcgacactgtcatgtgat 4908  
OY 3477 cactcacaagcctctcgtcaatcagaacgcagatttgatagtggtctgcacaatggaa 3536  
Db 4909 tgcacacgcctgtccacacatccagaatcagacttaatagtggttccagaaatggagc 4968

OY 3537 gataaagaacaaggaaactcatcagaagctccttgagaatcagacatatatttaagt 3596  
Db 4969 agtcaaggagatcgtgcacgcgtcagcagcgtcgtgcacagaagaagcattatttcaat 5028  
OY 3597 agtgaatgcacagtcag 3613  
Db 5029 gtcagtgctcagagctg 5045

RESULT 14  
AAZ49333 standard, cDNA, 3860 BP.  
XX  
AC AAZ49333;  
XX  
DT 14-MAR-2000 (first entry)  
XX  
DE Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.  
XX  
KW Multidrug resistance; MDR-1; P-glycoprotein;  
KW Transmembrane efflux pump; haematopoietic stem cell; transduction;  
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
KW gene therapy; gene replacement; genetic defect; thalassemia;  
KW Gaucher's disease; sickle cell anemia; leukemia; ex vivo expansion;  
KW cytokine; mutant; ds.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..3843  
FT CD5 /\*tag= a  
FT /product= "Human G185V mutant MDR-1 protein"  
XX  
XX W09961589-A2.  
XX  
XX 02-DEC-1999.  
XX  
XX 27-MAY-1999; 99WO-US11825.  
XX  
XX 28-MAY-1998; 98US-0086988.  
XX  
XX (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
XX Sorrentino B, Bunting K;  
XX  
XX WPI: 2000-072615/06.  
XX  
XX P-PsDB: AAY58187.  
XX  
XX Ex vivo expansion of hematopoietic stem cells transduced with a  
PT sequence encoding human multidrug resistance-1, used for bone marrow  
PT transplantation -  
XX  
XX  
XX  
XX  
XX Example 1; Page 79-82; 113pp; English.  
XX  
XX This sequence represents cDNA encoding human G185V mutant multidrug  
CC resistance protein MDR-1, where the the Gly residue at position 185  
CC of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a  
CC transmembrane efflux pump, responsible for the export of drugs from  
CC cells, particularly cancer cells. The wild-type MDR-1 shows increased  
CC resistance to etoposide and decreased resistance to vinca alkaloids  
CC compared with the G185V mutant. The invention relates to transducing  
CC hematopoietic stem cells with nucleic acid encoding an MDR protein  
CC and culturing the modified cells. The modified hematopoietic stem  
CC cells are useful in bone marrow transplantation (to reconstitute  
CC haematopoietic systems in patients who have undergone chemotherapy or  
CC radiation therapy) and in ex vivo gene therapy of genetic defects in  
CC cells derived from haematopoietic stem cells, e.g., thalassemia,  
CC Gaucher's disease, sickle cell anemia or leukemia. The modified  
CC cells can also be used to identify factors involved in regulating  
CC proliferation and differentiation in hematopoietic stem cells.  
CC Haematopoietic stem cells that express MDR-1 will be protected against

chemotherapeutic agents, so can be engrafted while the patient is undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells provides sufficient cells to permit standard biochemical analysis. Overexpression of MDR-1 allows cytokine-driven expansion of haematopoietic stem cells by at least 10-fold compared with a maximum of 4-fold in known procedures.

Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Query Match	36.3%	Score 1313;	DB 21;	Length 3860;
Best Local Similarly	62.2%	Pred. No. 0;		
Matches 2176; Conservative	0;	Mismatches 1270;	Indels 51;	Gaps 5;

OY	162	cagttgaccccttatattatgttggaaatagtgttcgccttgatttttgtttacaata	221
Db	336		395
OY	222	gatttccttgatataactcgaacagaacaggaaagattctgaanaacatttt	281
Db	396		455
OY	282	tcaatcagtttggcacaggacatcgctgctgttgatagctgtgatccgtgtaactaa	341
Db	456		515
OY	342	caatcgatatagacagattgac---aaatcagtaatgtatgtgaataagattgacct	398
Db	516		575
OY	399	gtttttcaaacaatgtctaactttttcgaatttgcctgtagctgtgtttgttgaagygct	458
Db	576		635
OY	459	gaaactaaccttagtactatcacagcgcctcctttaatgygcttaagcgagacgt	518
Db	636		695
OY	519	tctcagaatgtcatccatcttgacatgacagaaatlaagtctaatccaaaagctgggc	578
Db	696		755
OY	579	tgtgccaagaagaagctctgtgccatccagaaagfcatagcctttaaggcccagagaa	638
Db	756		815
OY	639	agaactcaaagratcacagatcatccaagaatgcaaaagatttggcataaaaaagac	698
Db	816		875
OY	699	tatagcttcaaaagtgtctctgtgtgctgtgtatcttctaagatggaaactatagact	758
Db	876		935
OY	759	tgcttttgttatgaaaccctcttgatttttaatgtagagacttgatatacatctgggac	818
Db	936		989
OY	819	tgtcttgcgttttctctttagtgttaatccataagcaatttgcattggagacagacatcc	878
Db	990		1049
OY	879	tcaacttgaacacttcgcaatagcccgagagagctgccttcaatbatlccagattatta	938
Db	1050		1109
OY	939	taagaacccccgatagataacttttccaaacttccaaagtgtatataaacctgaaatcagaag	998
Db	1110		1165
OY	999	aactgtggaatttaaanaatgttcttccaattatccatacaagaacatctatcaagattct	1058
Db	1170		1229

[illegible]

OY	2097	ttcatgtcagggatattttaa	tcg	gcagagcaggggaattt	taacga	tgaagatlaagaca	2156	
Db	2310	tttcttcaaggttttca	cat	tttgcgaagcttgga	gagatcttca	caacgcggtctcgata	2369	
OY	2157	cttgcccttcaagcca	atgttata	tcaggaat	tgcctgtgttga	gaaaggaacag	2216	
Db	2370	catgtttctcgatcca	atgctcaga	agagatgt	gagttgttga	tgacctaanaaac	2429	
OY	2217	caacagagcttgcaca	acatala	ttacga	taagata	tacgaacattcaagagcaacag	2276	
Db	2430	caatgagagatttgac	taccag	cgctcgcca	atgatgtctgc	caagttaaagggctatag	2489	
OY	2277	ttccaggaatggtcgt	cttaacaca	aaatgc	aaatcag	tgtggaacttccagtatacttc	2336	
Db	2490	ttccaggtcttgctga	atattacc	ccagata	tacgaacat	cttvgacaggaataattatc	2549	
OY	2337	ctttatatgtagtgga	atgatga	catcttc	gtatctc	gagatgtccagttacttgcgt	2396	
Db	2550	cttcatctatgtgtgc	aaactaac	taacgttacc	ctttagcaat	gtgtacccaatcttgc	2609	
OY	2397	gacagaatgatattga	aacccgcag	caatgtga	cttgatttgc	caacaagaatgaagaagact	2456	
Db	2610	agcagagagtgttga	aatatga	aaatgt	gtctgcga	acgaactgaagaagataagaagaact	2669	
OY	2457	taagcatctgtgaa	aaatagatgc	aaactgc	taaggaatata	cgtactatagtgatc	2516	
Db	2670	agaaggtgtctgtg	aaatctgc	tactta	gaacataga	aaacttccgaacggttgttcttc	2729	
OY	2517	aacaaaggagaaag	cgcttcg	agaatagt	atagaagag	gtcttcagaccacacagaa	2576	
Db	2730	gaactcagagcaga	agatttga	acatatgt	atgtctc	agagttgcaggttaccatacagaa	2789	
OY	2577	tacctcgaagaaag	ccagata	tattg	aaagctgt	tatgcatacgaatgccttata	2636	
Db	2790	cccttgaggaag	acacacat	ctttgga	attatactt	cttccacacgaagcaatgtatga	2849	
OY	2637	ttttgcctatgcag	cgaggttt	ccgatt	tgagcctat	ttaattcaagctvgagacgaatgac	2696	
Db	2850	tttttctctatgtc	gtatgttt	ctcggttttgc	agcctactt	gtgtgcataataactctag	2909	
OY	2697	cccagaagcgatgt	tcaatag	attttct	ccaattgc	ataagagctatgscatgtgaa	2756	
Db	2910	ctttgagagctc	gtgttga	tttccag	ctgttgc	cttcttvgatgcatgtgcgtvggga	2969	
OY	2757	aaacgcctgttcg	ccgcctga	atctccaa	gcgaacat	cggggctgcgagcatgtttgc	2816	
Db	2970	agtcagttcat	ttgtcc	ctcagact	atgc	aaagccaataatccagccacatcatcat	3029	
OY	2817	cttgtttgaa	aaagaa	ccaata	tagacag	cgcagtlcaagaaagggaanaagccagacac	2876	
Db	3030	gatcatctgaa	aaaaa	ccccctt	gtttgac	agctacagacaggaagcctaa	3089	
OY	2877	atgtgaaaggaat	tttag	gttttc	gagaag	ctcttcttcaatccatgtccagatgt	2936	
Db	3090	atgtgaaaggaat	aaatca	catctt	gtgaag	ttgtatcaactatccacccgaccagatcat	3149	
OY	2937	tttcaatcccg	tggtctta	ctccctc	agtat	ttvgcgaagaaagcaagtaacattgtg	2996	
Db	3150	cccagtgcttcag	ggagcttga	cgcttc	gtgaag	tgaagaaaggccagacgctgtcgtgtg	3209	
OY	2997	gagcagcgtgtg	ggaaag	caactc	gttca	actctgcagagac	cttataga	3056

Db	3390	catgtccctatggaagacaacgcccgggtgtgtctacagagaagatccgtggaaggcagcaca	3449
Qy	3227	tgccgcaatatccatctcttttatgaaagtccctccttgagaatatcaacacaaagtttgg	3296
Db	3450	ggaagcccaactaatactccttcatctgagtcactcgtcctaataatatactacacaaagtgg	3509
Qy	3297	actgaaggagcacagctcttcctgtgcgcgcagaacaagaactagctatctgcagaaggtctc	3356
Db	3510	agacaaaggaactcactgctctctgtgtgcgcagaacaacgcatctgcataagctctgcacct	3569
Qy	3357	tctccaaaaccccaaaatttatctgtgtgataggcgcacttaagcctcctcgataatagaag	3416
Db	3570	tgttagaagcgtccatatacttctgtcttctgtgataggaagccagtcgaactctgtgatacagaag	3629
Qy	3417	tgagaaggtgtgtcttagcatgctcctctgataaagccaggaagcggaagaaatgctctagtgt	3476
Db	3650	tgaaaaggtcttcccaagaagccctctgacaaagccagagaagccgcacactgtatgtat	3699
Qy	3477	cactcaacgggtctctctgcaattctcagaagcagaaattgataagtggtcttcgacaatatgaa	3536
Db	3650	tgtccacgcctgtgtccacatccagatccagaatcgaacttaaatagtggttctcgaatgtgcag	3749
Qy	3537	gataaaggaacaagaagcatctcaagaagctccggaagaaatcgagacatatatttaagt	3596
Db	3750	agtcagaagatgtgcacgcatcaagcagctgcctgcgcagacaagaagcatctattttcaat	3809
Qy	3597	agtgaaatgcacagtcag	3613
Db	3810	ggtcagttgtccaagctg	3826

RESULT 15  
AAZ24041  
ID AAZ24041 standard; cDNA; 8630 BP.  
XX  
AC AAZ24041;  
XX  
DT 04-FEB-2000 (first entry)  
XX  
DE Retroviral M4 mdr-1 cDNA.  
XX  
KW Retroviral vector; gag gene; gene therapy; chemotherapeutic agent;  
KM hematopoietic stem cell transformation; mdr-1; ss.  
XX  
OS Retrovirus.  
XX  
PN EP955374-A2.  
XX  
PD 10-NOV-1999.  
XX  
PF 07-MAY-1999; 99EP-0250151.  
XX  
PR 08-MAY-1998; 98DE-1022115.  
XX  
PA (PETT-) PETTE INST HEINRICH.  
XX  
PI Ostertag W, Baum C, Hildinger M;  
XX  
DR WPI: 2000-001087/01.  
XX  
PT New retroviral vector containing minimal or no gag gene sequence, for  
PP use e.g. in gene therapy or cloning -  
XX  
PS Disclosure; Page 16-18; 35pp; German.

This invention describes a novel retroviral vector (RV) containing a gag  
CC gene fragment having fewer than 400 bp. RV are used: (i) in gene therapy  
CC ((ii) for cloning genes; (iii) for (over) expression of proteins or RNAs  
CC and (iv) for transfection of hematopoietic stem cells (especially to  
CC impart resistance to chemotherapeutic agents). Because of their reduced  
CC content of viral genes, RV are very safe (no expression of toxic or  
CC immunogenic proteins, no recombination with other viruses), have







---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:50:48 ; Search time 14771.4 Seconds

(without alignments)  
4044.048 Million cell updates/sec

Title: US-09-873-409-14

Perfect score: 3621

Sequence: 1 ttccgcttgctgctgact.....atgcacagctcagctgta 3621

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2344280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml:\*

- 1: gb\_ba:\*
- 2: gb\_dtg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_om:\*
- 20: em\_or:\*
- 21: em\_ov:\*
- 22: em\_pat:\*
- 23: em\_ph:\*
- 24: em\_pl:\*
- 25: em\_ro:\*
- 26: em\_sts:\*
- 27: em\_sy:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htgo\_hum:\*
- 31: em\_htgo\_inv:\*
- 32: em\_htgo\_rod:\*
- 33: em\_htg\_hum:\*
- 34: em\_htg\_inv:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1325.4	36.6	4279	6	AX105078 Sequence
2	1325.4	36.6	4279	6	AX105080 Sequence
3	1325.4	36.6	4279	6	AX105082 Sequence
4	1323.8	36.6	4279	6	AX105057 Sequence
5	1322.2	36.5	4045	12	AF269224 Synthetic
6	1320.6	36.5	4317	4	AF045016 Canis fam
7	1320.6	36.5	4317	6	AX105059 Sequence
8	1316.2	36.3	4646	6	I49610 Sequence 2
9	1316.2	36.3	4646	6	HUMMDR1
10	1316.2	36.3	4669	6	I08557 Sequence 3
11	1316.2	36.3	6505	6	AR028671 Sequence
12	1316.2	36.3	9318	6	AR028672 Sequence
13	1314.6	36.3	4378	6	E02326 Multidrug r
14	1314.6	36.3	8630	6	AX012321 Sequence
15	1311.4	36.2	8630	6	AX012320 Sequence
16	1308.2	36.1	4669	6	AR055785 Sequence
17	1308.2	36.1	4669	6	AR091275 Sequence
18	1306.6	36.1	4186	6	AX108654 Sequence
19	1306.6	36.1	4195	6	AX108656 Sequence
20	1305.2	36.0	4264	6	AR051647 Sequence
21	1303.6	36.0	4264	6	AR051650 Sequence
22	1300.2	35.9	3988	6	AX024454 Sequence
23	1300.2	35.9	4192	9	AF016535 Homo sapi
24	1280.2	35.4	4018	5	GGA9799 Gallus ga
25	1276.2	35.2	3924	6	AX024455 Sequence
26	1276.2	35.2	3924	9	HUMMDR3
27	1276.2	35.2	4296	10	CRUPGPI
28	1274.6	35.2	3987	10	CRUPGPI185
29	1274.6	35.2	4304	10	CRUPGPI165
30	1272.8	35.2	4280	10	CRUPGPI1
31	1253.8	34.6	3858	4	OA078609 Ovis aries
32	1218.6	33.7	4084	10	MUSMDRA
33	1218.4	33.6	4927	10	AF257746 Rattus no
34	1216.4	33.6	4298	10	MUSMDR
35	1215.8	33.6	4323	10	AF286167 Rattus no
36	1212	33.5	4233	6	AR123273 Sequence
37	1209	33.4	3905	10	CRUPGPI11
38	1204.2	33.3	4924	4	MUSMDR1A
39	1202.8	33.2	3489	4	AB029153 Felle cat
40	1202.6	33.2	4356	10	MUSMDRAA
41	1184.8	32.7	4254	10	RATMDRM
42	1184.6	32.7	4390	5	XU017608 Xenopus lae
43	1180	31.5	3912	10	RATPGLYCO
44	1139.4	23.9	3682	10	MUSPGLYXA
45	865	23.9	4280	9	AF136523 Homo sapi

## ALIGNMENTS

RESULT 1  
LOCUS AX105078 4279 bp DNA  
DEFINITION Sequence 22 from Patent WO0123540.  
ACCESSION AX105078  
VERSION AX105078.1 GI:13921228  
KEYWORDS  
SOURCE  
ORGANISM  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 4279)  
Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Reif,T.C. and  
Patten,C.J.  
P-glycoproteins and uses thereof  
Patent: WO 0123540-A 22 05-APR-2001;  
GENTEST CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..4279  
/organism="Canis familiaris"

TITLE  
JOURNAL  
FEATURES  
source

CDS		/db_xref="taxon:9615"	
		17..3862	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAC37721.1"	
		/db_xref="GI:13921229"	
		/translation="MDPEGGRKGSAEKFMKMKSKNEKKKPYSTFAMFRYSN	
		WLDRLVLMVGTMAALIHGAALPLMLVFGNMTDSTANAGISNKTFPIINISNTNT	
		OHFNIHLEEMETIAYVYSGIGAGVLAAYIOVSWPCALAGQILIKIRQEPHLMRQ	
		EIGMDVDEVEELNTRFDVSKINEGIDGKIFGPFQIATFEPFIATFEGFWMLTL	
		VLIATSPVIGSAIAWAKITLISFTDKELIYAKAGAAVEVLAIRTYAIFSGOKEL	
		EYRNKLEPAGIGIKRAITANISIGAFILIAVSAVALAFVSTVLSSEKIOVL	
		TVFISVLGARSIGASPSIEFANARGAATFEITIDNKPESIDYSKSGHKPDNKG	
		NLEFKNVHSYPSRREKILKGLNLKVSQGVQALVANGSGCKSTVDLMQLPDPT	
		GMVCIIDGODIRITINVRHLREITVYQSOEVLATIAERIGRENVNTELEKAVE	
		ANAYDFIMKLPKPEFTLVGERGASLGSGQALKAIRALVRNKLILLDEASADTE	
		SEAVVOVALDKARKGRTTIVIAHRLSTYRNADVIAGFDGVVPEKNHDELMEKGIY	
		EKVIYMOIRGNIELENNATGESKSDALEMSPKSGSLIRSRTRRTHAPQODR	
		KLGTEDLNENPVSVFNLKLNSTENPVPVGVJFCALINGLOPAPESIIISRIIGT	
		FTRDDEPETERONSMFVSLFLVGIISFTFFLQFTFGKAGELLTKRLRVMFRSM	
		LRQDVSWFDDPKNTGALTTRLANDAQAQVKAIGSLAVITONINLGTGIIISLIG	
		WOLLLIATPIIAGVVEKMLSGQALDKDKELEGAKLATEIENFRVIVSLTR	
		EOKFEYVAQSLQVPRNSLRKAHIEGVSESTOAMVPSYAGCPREFGAYLVANEFM	
		FQDVLIVSAIFYGAMAYQVSSFAPDYAKAKVSAAHVIMLIEKSPILDSYSPHGIKP	
		NLEBSNVTFNEVFNPTIRPDIPLVLOGSLVETKKGQTLALVSSSGCKSTVOLLEF	
		YDPLAGSVLIDKEIKHLNVQWLRALHGLIVSQEILFDCSIAENLAVGNSRVSHEE	
		IMQAARKANIHFIETLPEKYNTRVGDKGTQISGGQKQIARALVRPPLLDEA	
		TSADTSEKRYVOEALDKARKEGRTCIIVIAHRLSTIONADLIIVFONGKVEKHQOL	
		LAQKGIYSMVSVOAGAKR"	
BASE COUNT		1295 a	833 c 1008 g 1143 t
ORIGIN			
Query Match		36.6%; Score 1325.4; DB 6; Length 4279;	
Best Local Similarity		62.6%; Pred. No. 3.7e-268;	
Matches 2190; Conservative 0; Mismatches 1256; Indels 51; Gaps 6;			
Qy	162	caagttgacccgtatataatgltggaataagtgltgctgcctgaatttttggttaacata	221
Db	358	CACGTAATGCTTATTATACAGTGGGATCGGCTGGCTGCTGCTGCTTACATCA	417
Qy	222	gattctctgtgataataactgcaagcagcagcacaagaagatttcaaaacgctttc	281
Db	418	GGTTTATTCTGTGCTCGCAGACGAGAAAGACAGATCTCAAAATTAAGAAACATTTT	477
Qy	282	tcattcaagtttgcaagcagacatcgctgttgataagctgtgacatcggtgaactaa	341
Db	478	TCATGCTATCATGCGACAGAGAGATGGCTGTTGAGTGCATGACGTTGGGAGCTTA	537
Qy	342	cactgcgatgacagacatg--acaaatcagtgatgltatggaataagattgctct	398
Db	538	CACCCGGCTCACAGACAGATGCTCCAAATCAATGAAGGAATTGGCAAAATTTGGAAT	597
Qy	399	gtgtgttcaaaacatgtctactcttttcgatgtgcctgagctgtgtgtgtgtaagagctg	458
Db	598	GTTCTTTCAATCAATACCAACATTTTACCCGGTTTATATAGTGGGTTTACACGCTGTTG	657
Qy	459	gaaactaacctagtgatctatccacgctctcctctataatgcttcaagcggcagcatg	518
Db	658	GAAAGCTAACCTTGTGATTTTGGCCATCAGCCCTGTTCTTGACATTCAGCCGCCATCTG	717
Qy	519	ttctaagaatgltcatctatgacaglaaggaatlaagltgcatttccaagctggggc	578
Db	718	GGCAAAATATCTATCTCATTTACTGATTAAGAAACTCTTGCCCTATGCAGAAAGCTGAGC	777
Qy	579	ctgtggcagaagaagctctgtcatcatccgaacgataagcctttagggcccgagaa	638
Db	778	AGTAGCTGAAGAAAGCTTTAGACAGCACTGTGATTCCTTTGGAGGACAAAGAA	837
Qy	639	agaactcaagaagtatacaagaatccaagaatccaagaagatttggcataaaagaagc	698
Db	838	AGAAGCTTGAAGAGTACAAATAATTTAGAGAAGCTTAAGGAATTGGGATTAAGAAAGC	897
Qy	699	tatagttcaaaaagtgtctctgtgtgctgtgatactcttattatgaatggaaacctatgact	758
Db	898	TATCAGCGCCACCATTTCTATTGTGGCCGCTTCTTATTGATCATGACATCATGCTCT	957
Qy	759	tgcttttggtaagaaacctcctctgtattcttaatggagaacctgatatcaactcgggac	818
Db	958	GCGTTTCTGTGTATGGGACCTCCTTGCTCTCCAG-----TGAAATATCTATTGGACA	1011
Qy	819	tgltctgtcttcttcttcttagtgttaatccatagcagttatctgattggagcagcagttccc	878
Db	1012	AGTACTACAGCTCTCTCTTTCTATTAAATTTGGGCTTTTATGTATGGACAGGCATCC	1071
Qy	879	tcacttgaaccttcgccaatagcccgagagctgtccttcatatttccaggatttga	938
Db	1072	AAGCATTTGAAGCATTTTGCAAAACCCAGAGAGACAGCTTATGAATCTTCAAGTAATTTGA	1131
Qy	939	taagaaaccaglatagataacttttccacagctgatatlaaacctgaaatccatagaagg	998
Db	1132	CAATTAACCAAGCATTTGACAGCTATTGAGAGTGGACATTAAGCAAGATTAATTAAGCG	1191
Qy	999	aactgtgaaatttaaaaatgttcttccaattatccatcaagaccatctatcaagattct	1058
Db	1192	AAATTTGGAAATCAAAATATGTTCACTTCAAGTTACCCCTTCGAAAAGATTAAGATCTT	1251
Qy	1059	gaaaggtcgaatcctcgaatlaagctcgtgagagacagtcgcctgtgctcctcaatgg	1118
Db	1252	AAAGGGCTCAACCTTGAAAGCTTCAAGATGAGATGGGACAGCTGGCTGTTGGGAACTGG	1311
Qy	1119	cagtggaagaagtaagcgtatgccagctcttcgcaagagttatataatccgatgatgctt	1178
Db	1312	CTGGGGGAAGACGACGACCGGTGATGACAGAGCTGATGACAGGCTGATACCCACAGATGGCAT	1371
Qy	1179	tatcatggtgagatgagataatgacatcagagctttaaagtgcgagattatagagccat	1238
Db	1372	GGTCTGATTGATGACAGACGACATTAAGACCATTAAGATTAAGGATCTTGGGAAATTAAC	1431
Qy	1239	tgagtggttgaagcaagagcctgttctgttcgagaccacatccatgataacatcaagta	1298
Db	1432	TGGTGTGTGATGATGAGAGAGCTGTGTTGTTGCCACACGATTAAGAAATTTCCGCTA	1491
Qy	1299	tgagcagatgatgtgactgataagagaatgagagaagcagaaggaaggaacaaatgcgtla	1358
Db	1492	TGGCGGGAAATATGTCACCATGATGAGATGAGATTAAGAAAGCTGTTAAGGACCAATGGCTTA	1551
Qy	1359	tgattttatcatgtatgttctctaataatttaacatctgtgtaagggaaagaagagctca	1418
Db	1552	TGATTTATCATGGAACCTACCTTAATAATTTGACACTGTGGTGGAGAGAGAGGGCCCA	1611
Qy	1419	aatgaatggagggcagaagaacagagatgcgaattgctgccttaagtctcgaaccccaa	1478
Db	1612	GCTGAGTGTGTGACAGAAACAGAAATGCGCATTTGCTGGGCCCTGTTGGCAACCCCA	1671
Qy	1479	gattcgtatttagatagagctgctgcctctgattcagaagaagaatcagctggttca	1538
Db	1672	GATTTCTTGTGTGATGAGGACAGACGCTGCTGGACCTGAAGTGAAGCAGTGGTTC	1731
Qy	1539	agctgcactggaagaagcagcagaagatcgactacaaatcgttgtagaacaccgacttc	1598
Db	1732	GGTGCCCTGTGATTAAGGCCCAAAAAGGCCGACACCACTTGTGATAGCTCATCGTTTTC	1791
Qy	1599	tactattgaaagtgcagattgattgtgtgacccctaaagatggaatgcggcgagaaagg	1658
Db	1792	TACAGTTGTGTAATGCGCGATGATTTGCTGTTTGTGATGTGAGATGATTTGGGAAAGG	1851
Qy	1659	agcaatcgtgaactaaatggaanaaagcagatctatattatcaactctgtgattgcacagga	1718
Db	1852	AAATCATATGATCACTGAAAGAGAAAGAGAGGCACTTTACTTCAAACTTGTCAACATGACAC	1911
Qy	1719	tattaaaaaagctgata--acagatggagatcaatgacatattctactgaaagaagacaa	1777
Db	1912	AAGAGAAATGAATGATTGATTAAGAAATGCACTGCTGTAATCCAAATGTGAAGATGATGC	1971
Qy	1778	actcaatccctcgtcacctcgtgtaagag-----catcaagtcagactcttcatgacaa	1829

Db 1972 CTGGAATGTCCTCCAAAGATTTACAGGCTCCAGTTTAATATAAAGAGATCAACTGCAG 2031  
 QY 1830 ggcgtagaagatccaccaccaacttaagagataagctt----- 1866  
 Db 2032 GAGTATACATGACCAAGGCCAGACAGAAAGCTTGGTACAAAAGAGACTTGAATGA 2091  
 QY 1867 -----cctgaagctcctcttaaaaaatttaagtaaaacagcctgaatgacctt 1919  
 Db 2092 GAATGTACCTCCAGTTTCCCTTGAGGATTTCTGAAGCTGAAGCTCAAGTGAATGGCTTA 2151  
 QY 1920 tctgtctcgtgagacatctgctctcttcaaatgaaactgtctcatcagatatttccat 1979  
 Db 2152 TTTTGTGGTGGTATATTTTGGCTATTTAATACGAGGCCCTGCACACGACTTTTCAT 2211  
 QY 1980 catcttgcnaaaattataacacatgltt--gnaataaagtataaaaccacatlaaaga 2036  
 Db 2212 AATATTTTCAAGATTAATAGGATCTTTACCCGAGATGAGATCCTGAATAAAGACGACA 2271  
 QY 2037 tgaicagaataattatccatgatattgcatattggatggtatattgcttgctgaatta 2096  
 Db 2272 GAATGATACATGTTTCTGATTTGTTCTAGTCTCTGSAATTAATTTCTTTATTAATCATT 2331  
 QY 2097 ttccatcgaggaattatttaacgcgcagcagcagcagcagcagcagcagcagcagcagc 2156  
 Db 2332 TTTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCCTCATAACGGCTTGATGA 2391  
 QY 2157 cttgacctcaaacgaatgtatatacagaatattgacctggttgaatgaagaagaacag 2216  
 Db 2392 CATGTTTTCAGATCCATGCTGAGACAGAGATGCTAGCTGTTGATGACCTTAATAAACAC 2451  
 QY 2217 caccaggaagcttgacaataatattagccatagatatagacaataatcagcagcagcag 2276  
 Db 2432 CACTGGAGCATTTGACCAACAGGCTTGCCATGATGCGCTCAAGTTAAAGGGCTATTAG 2511  
 QY 2277 ttccaggaattgctccttaacacaaatgcaactaacaactgagcacttccatattc 2336  
 Db 2512 TTCCAGGGCTGCTGATTAATACCCAGATTAAGCAATCTTGGAGACGGATTAATTAATC 2571  
 QY 2337 ctttatataatgagtgagagatgacatctcctgattcctgagatattgctcagactg 2396  
 Db 2572 CTATATCATGCTTGGCAATTAACACTTTTACTTACCAATGTTGCCATCATTTGCAT 2631  
 QY 2397 gaccagaatgatattgaacccgcagcagcagcagcagcagcagcagcagcagcagc 2456  
 Db 2632 AGCAGAGAGTGTGTAATGAATAATGTTGTCGACACAGCAGCAAGAAAGATGAAGAGCT 2691  
 QY 2457 taagcatgctgaaagatagcaactgaaagccttgagagaaatatacgtactatagtc 2516  
 Db 2692 AGAAGAGCTGGGAAGATTTGCTACAGAACCATCGAAACTTCGCACTGTTTCTTT 2751  
 QY 2517 aacaaaggaaaaagccttcgcagcaaatglatgaagagatgcttcagactcaacacaga 2576  
 Db 2752 GACTGGGGAGCAGAAATTGAATACATGATGACAGAGTTTGACAGTACCATACAGAA 2811  
 QY 2577 taacctgaagaagaacagattattggaagctglatatgcatccagcagcagcagcagc 2636  
 Db 2812 CTCTTTGAGGAAGACACATCTTCGGGGCTCATTTTCTATACACCAGCAATGATGA 2871  
 QY 2637 ttctgcatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2696  
 Db 2872 TTTTTCATGCTGGCTGTTTCCGGTTTGCTGCTACTGTTGGGCAAAATGAGTCAATGA 2931  
 QY 2697 cccagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2756  
 Db 2932 CTCTTACGATGTTCTTTTGGTATTCACACTATTGCTTTGGGACATGGCAGTGGGCA 2991  
 QY 2757 aacgctgcttggctcctgaatattccaaaccaaactcggggcctgcgcagcagcagc 2816  
 Db 2992 GGTCAATTCATTTGCTGCTGACTATGCAACCAAGATATGACGACCCAGCTCATCAT 3051  
 QY 2817 ctgtctgaaagaaacaaatatagacagcgcagctcaagaggaagaaagccagacac 2876

Db 3052 GATCATTTAAAAAGCCCTCTGATGACAGCTACAGCCCTCACGGCCTCAAGCCAAATAC 3111  
 QY 2877 atgtagaaggaaatttagagtttcgagaagctctcttcttcatccatgltgcagagatg 2936  
 Db 3112 GTTGAAGGAAGAAAGTACATTTTAATGAGCTGTTTCACATATCCCATCTGACAGCAT 3171  
 QY 2937 ttcatctccctggtcttccctcagatattgagcaggaagaacagtaacatgcttgagg 2996  
 Db 3172 CCCCCTGCTCCAGGGCTGAGCTCGAGTGAAGAAAGGCGCAACGCTGACCTCGTAGG 3231  
 QY 2997 gacgaagcgtctggtggaagaaacactctcttcaactctgcaagacattatgaccctg 3056  
 Db 3232 TACAGATGGCTGTGGGAAGACAGATGTTGCTGACCTCTAGACGGCTTATGACCCCT 3291  
 QY 3057 gcaagcaagctgctgttgatgagtgatgagcgaagaatgaatgagtagagctgcg 3116  
 Db 3292 GGTGCTGATGCTAATGATGAGCAAGAGATTAAGACACTTAATGCTCAGTGGCTCCG 3351  
 QY 3117 ttccaaatagcaatgcttcccaagcgtgctcttcaactctgcaagcagcagcagcagc 3176  
 Db 3352 AGCACACCTGGGATGCTGCTCAGAGACCCATCTGTTGACTGACGATTCGCCAGAA 3411  
 QY 3177 catgcctatggtgacaacagccgtgltggtccatagatgagatcaagaagccgcaaa 3236  
 Db 3412 CATTTGCTATGAGACACAGCCGGGCTGATCATGAAGAGATTAATGACAGCAGCA 3471  
 QY 3237 tgcagcaaatatcatctcttattgaagctcccttgagaaataaacaacaaagtgtg 3296  
 Db 3472 GGAGGCCAATACACACTTCATCGAGACACCTCCGAAATTAACACACAGAGTAGG 3531  
 QY 3297 actgaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3356  
 Db 3532 AGCAAAAGACACCCAGCTCTGCTGTGGCCAGAAACAGCAGCTTGCATGCTGCGCT 3591  
 QY 3357 tctccaaaaccccaaaatttattgltgagtgagcagcagcagcagcagcagcagcagc 3416  
 Db 3592 TGTATACAGCCCTCATATTGCTTTGATGAAGTACATGAGCTGATGATGAGAAAG 3651  
 QY 3417 tgagaagctggttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3476  
 Db 3652 TGAAGAGCTGTTCCAGAAAGCCCTGAGCAAGAGCAGAGAGGCGGCACTGTTGAT 3711  
 QY 3477 cactcagagcgtctctgcaatcagaaagcagagcttgatgagtggtctgcagacagtaa 3536  
 Db 3712 CGCCACCGCTTCCACACCTCCGAATGCAATTAATGATGATGTTTCAAGATGCA 3771  
 QY 3537 gataaaggaacgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3596  
 Db 3772 AGTCAGAGGACATGGCAGCATACAGCTGCTGGGCCCAAAAGGCAATCATTTTCCAT 3831  
 QY 3597 agtgaatgcagagtcag 3613  
 Db 3832 GGTCACTGTCACAGCTG 3848

RESULT 2  
 AX105080 4279 bp DNA PAT . 30-APR-2001  
 LOCUS AX105080  
 DEFINITION Sequence 24 from Patent WO0123540.  
 ACCESSION AX105080  
 VERSION AX105080.1 GI:13921230  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 1 (bases 1 to 4279)  
 STOCKER,P.J., Steinel-Crespi, D.T., Crespi, C.L., Relf, T.C. and  
 Patten, C.J.  
 P-glycoproteins and uses thereof  
 Patent: WO 0123540-A 24 05-APR-2001;  
 JOURNAL GENTEST CORPORATION (US)  
 FEATURES Location/Qualifiers





OY	1778	accactctccctcgcactctcgtgaaag-----calcaagtaagactcattgaca	1829
Db	1972	CTTGGAAATGTCTCCCAAAAGATTCAAGGGTCCAGTTTAAATAAAGAAATCACTCGCAG	2031
OY	1830	gacgtgaagatccaccacaaactaaagaagaagtcctt-----	1866
Db	2032	GAGTATACATGCAACCAAGGCCAACAACAGAAAGCTTGGTACAAAGAGACTTGAATGA	2091
OY	1867	-----cctgaagtcctctatataaanaattlaaagttaacaagccctgaatgacctt	1919
Db	2092	GAATGTACTCTCAAGTTTCCCTTGTGGAGATTCCTGAAGCTGAACCTCACTGAATGGCCTTA	2151
OY	1920	tgtgtctcgtgggacatggtcctcgtctctaaaygaactcgtlccatcagatltccat	1979
Db	2152	TTTTGTGTTGGTGAATATTTTGTGCTATATTAACCGAGGCCCTGCAACCAAGATTTCAT	2211
OY	1980	catcttgcgaanaattataacccgtt---ggaataagtgtaaaccacattaaagca	2036
Db	2212	AATATTTTCAAGGATTAATAGGATCTTTACCCGAGTGGAGATCCTGAACCAAAACGACA	2271
OY	2037	tgaagcagaatltatccatgatalatcgcatlttggatgtatcttgcgtlccagta	2096
Db	2272	GAATAGTAACATGTTTCTGTATTTGTTTCTAGTCCCTGGAATATTCTTTATTACTTT	2331
OY	2097	ttccatgcaggagtalatttacgcgcagcagcagggganaatttaacgatyagattaa	2156
Db	2332	TTTTCTCCACAGGGCTTCAATTTTGGCAACGTGGGAGATCCTCAATAGCGGCTCGATA	2391
OY	2157	cttgccttcaaaagccatgtlatalatcaagatalatgcctggttggatgaaagaaacag	2216
Db	2392	CATGTTTTCAATCCATGCTGTGAGACAGATGTCACCTGTTGATGACCTTAATAAACAC	2451
OY	2217	caacggagccttgaacaacatctagcatalagatalagcaacatccacagcagcaagc	2276
Db	2452	CACGTGAGCATTGCAACCAACGCGCTTCCCAATCATGCGGCTCAAGTTAAAGGGCTATAGG	2511
OY	2277	ttccaggatctggcctcttaacacanaatgcaactaaacatcagtgacttcatlcatc	2336
Db	2512	TTCCAGGCTTGCTGTCTATTAACCCAGAAATTAAGCAAACTTTGGACAGGCAATTAATATATC	2571
OY	2337	cttataatagatggagagatgacatltcctgattcgtagatctgctcgaagtctgcgt	2386
Db	2572	CTTATCTATGTTGGCAATTAACACTTTTACTCTTAGCAATTTGACCACTCATCTTGCAAT	2631
OY	2397	gacagcagatgtgaanaacccgcagcagatgctgattgctcaacaagaagtaagcaagact	2456
Db	2632	AGCAAGAGCTTGTGAATTAATAATGTTGTGTGACAGCAAGCACTGAAAGATTAAGAAAGACT	2691
OY	2457	taagcatgctcgtgaagaagatgaacaactgaagccttggagaatalagtlactatagtcct	2516
Db	2692	AGAAAGAGGTGGGAAGATTCTCAAGAGCGAATCGAAAACTTCCGAACGTGTTCTTCTT	2751
OY	2517	aacaagggaanaagccttcgagcaaatgtatgaagaatgacttcagactcaacagagaa	2576
Db	2752	GACTCGGGAGCAGAAATTTGAATACATGTATGCACAGAGATTTGCAAGTACCATACAGAAA	2811
OY	2577	tactctgaagaagaacaaagttatctggaagcgttlatgcatccagcaatgccttatata	2636
Db	2812	CTCTTTTAGGAACACACACTCTTCGGGCTCATTTTCTATACCCACGCAATGATGTA	2871
OY	2637	tttgcctatgcacagcaggttctgattggagcctatcttaatcaagctcgtgacgtac	2696
Db	2932	CTTTCAGAGATCTCTTTGGTATTTCAAGATTGCTTTGGTGCACATGACAGTGGGCA	2991
OY	2757	aagcgtcgttctgctcctcgaatatcccaagcacaatcgcggggtcgtgcgcatctgttc	2816
Db	2992	GGTCAGTTCATTTGCTCTCTCAATAGCAAAAGCAAAAGTATCAGAGCCACGTCATCAT	3051

QY	2817	cttgcttgtaaaaaagaaccacaataatagacagccgcgaagtcacaagaagggaaaaagccagaac	2876
Db	3052	gatctattgtaaaaaaacccctctgattggaagctacagcccttcacagccctcaagccaaatpac	3111
QY	2877	atgtaaagggaatttagagctcttcagagaagctctctctctctctatccatgctgcgccagat	2936
Db	3112	gttggaaaggaatgtgacattttaattgaagtcgtgttcacatctccacactgcacacagacat	3171
QY	2937	ttccatccctcgttggtctatccctccatgattatgaagcgaagaaagacagtagcatltytgg	2996
Db	3172	ccccctgctccagagggcctgcagccttcagagtgtaaaagagggcagacagctggccctcgtgag	3231
QY	2997	gagcagcgagctgctggaagaaagcactctgctcaactctctgaaagacttatgacccct	3056
Db	3232	tagcagctggcgtgtgggaagacacagttgttcacctctctacagggccttataatgacccctt	3291
QY	3057	gcaaggaacaatgtcgtttgtatgtgtgtagtgcaaaagaatctgaatgtacagtgtcccg	3116
Db	3292	ggcttggttcagtgctcatttgatggcaaaagaaatpaaagcactgtaattgtccagtggtccg	3351
QY	3117	ttcccaaaatagcaatcgtcttcctcaagaagcctgtgctcttcaactgcagacttgctgaa	3176
Db	3352	agcacacccctgggacatcgtgtctcagagagccacatccctgtttactgcagacattggccagaa	3411
QY	3177	catgcctatggtgacaaacagccgtgtgtgtccatttagatgagatcaagaagccgcgcaa	3236
Db	3412	catgtccctatgtgagacaacacccgggtcgtatcacatgaaagatpattgtcagcgacgcaa	3471
QY	3237	tgcagcaaatccatctctttattatgaagtcctccctgagaataaacaacaagafttg	3296
Db	3472	ggagggccaacatcacacacacttcatcgagacacctccctgagaatpaaacacagagtagg	3531
QY	3297	actgaaagagacacagcctcttcgacgcgcgaagaaacaagactagctatltgcaaggtct	3356
Db	3532	agacaaagaaacccagcctctcgtggtggccaaanaagggcattgtccattagctgcgcgtct	3591
QY	3357	tctccaaaaccacaatttatgttgaatgaagccacttcagccctcgatlaatgacag	3416
Db	3592	tgtttgacacagcctctattttgtcttttgatgaagactpacatcacctctgattacgaaag	3651
QY	3417	tgaagaagttgttgaagctgcctctgataaagccaagacggaagagacatgcccagtggt	3476
Db	3652	tgaaaaggctgttcacaagaagcccttcgacaaagccacagaaagccgcaccttcatttgat	3711
QY	3477	caactcaagcgtctctgtcaattcagaagcagatttgaatgtgttctgcacaatgaaa	3536
Db	3712	cgccacacgcgttgttcaccatcacaatgcagaatttaattatgtgcttgcgaattggcaa	3771
QY	3537	gataaaggaaacaagaactcatcaagagctcctctgaaatctcagagacabatatttaagtt	3596
Db	3772	actcaaggagcatgtgcacacacatcacacctcgtggccagaaagacatctatttttccat	3831
QY	3597	agtgaatgcacagtcag	3613
Db	3832	ggctcagctgcacggctg	3848
RESULT 3			
AX105082			
LOCUS	AX105082	4279 bp	DNA
DEFINITION	Sequence 26 from Patent WO0123540.		
ACCESSION	AX105082		
VERSION	AX105082.1	GI:13921232	
KEYWORDS			
SOURCE			
ORGANISM	dog.		
	Canis familiaris		
	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;		
	Mammalia; Mammalia; Carnivora; Fissipedia; Canidae; Canis.		
	1 (bases 1 to 4279)		
	Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and		
	Patlen,C.J.		
	P-glycoproteins and uses thereof		
	Patent: WO 0123540-A 26 05-APR-2001.		
TITLE			
JOURNAL			



Db 1912 AAGAGAAATGAATGAGTTAGAAAATGCCACTGTGATTCACAAAAGTGAAGTATGAC 1971  
 Qy 1778 actcaatccctcgtgacctcgtgaaag-----catcaagtgaagcttcaatgacaa 1829  
 Db 1972 CTTGGAAATGCTCCAAAAGATTGAGGGTCCAGTTTAAATAAAGAAATCAACTACGACG 2031  
 Qy 1830 ggcgaaggaatcccaaccaactaaagagataagtcct----- 1866  
 Db 2032 GAGTATACATTCACCACACAGGCCAAGACAGAAAGCTTGTCACAAAAGAGACTTAAATGA 2091  
 Qy 1867 -----cttgaagtcctcattaaataatlttaaaglttaaagacctgaagtcctt 1919  
 Db 2092 GAATGTACCTCCAGTTTCTTCTGGAGATTCGTGAAGCTGAACCTCACTGAATGAGCCCTTA 2151  
 Qy 1920 tctggtctcgtgagacatcgtgctctcgtcctaaagaaagtcgtcaagtaatttccat 1979  
 Db 2152 TTTTGTGGTTGGTATATTTTGTCTATTATTAACGAGGCGCTGCACACGACTTTTCAAT 2211  
 Qy 1980 catcttgcacaaatataaacaatgctt---gaaataatgataaacaactaaagca 2036  
 Db 2212 AATATTTTCAAGGATTAATGAGATCTTACCCGAGATGAGATCCTGAACAAAACGACA 2271  
 Qy 2037 tgaatgcagaaattatccatgataatcgtcaatttgggtgltatltgcttgcagta 2096  
 Db 2272 GAATAGTAACATGTTTCTGATTGTCTGATGCTTGAATTAATTTCTTTTATTAACAT 2331  
 Qy 2097 ttcaatgcagagatatttccgcgcagagcaagggaatlttaacatgagatlaagaca 2156  
 Db 2332 TTTTCTCCAGGCTTTCACATTTTGGCAAAAGCTGGGAGATCTCCTAATGACGCGCTTCGATA 2391  
 Qy 2157 ctggccttcaagcaatgataatcagatataatgctggttggttgtaagaagaaagcaag 2216  
 Db 2392 CATGTTTTCAGATCAGCTGAGACAGAGATGACGTGGTTGATGACCTTAATAAACAC 2451  
 Qy 2217 cacaagagcttgcacaaatataatgacatagataatgacaaatcagagcaagcaag 2276  
 Db 2452 CACTGAGACATTTGACAAACACGAGCTGCCAATGATGCGGCTCAAGTTAAAGGGGCAATGAG 2511  
 Qy 2277 ttcaagatctggcgcttaacacaaatgcaactaactaagtggaacttcaatcatctc 2336  
 Db 2512 TTCCAGGCTTCTGATACCTACCCAGAAATATAGCAAAATCTTGAGACAGGCAATTAATATAC 2571  
 Qy 2337 cttatataatgagtgaggagatgaacttccgtatcttgatgtatgctccagtcactgcgct 2396  
 Db 2572 CTTTATCTATGTTGGCAATTAACCTTTTACTTACCAATTAATTAATTAATTAATTAATTAAT 2631  
 Qy 2397 gacaggaatgattgaaacccgcagcaatgagcttgcacaaagaaatgaagcaagcaat 2456  
 Db 2632 AGCAGAGATTTGTAATGAAAATGTGTCTGACACAGCAGTGAAGATGAAGAAAGAGCT 2691  
 Qy 2457 taagcatctggaagaatagcaactgaagcttggagaataacgtactatagtcact 2516  
 Db 2692 AGAAGGAGCTGGAGATGTTGCTACAGAAAGCCATCGAAAACCTCGAAGCTGTTTCTTCT 2751  
 Qy 2517 aacaaggaagaaagcctcgcgcagcaatgataagagatgctcagactcaacacagaa 2576  
 Db 2752 GACTCGGAGAGAAATTTGATATCATGTATGACAGAGTTTGGCAATGACATACAGAAA 2811  
 Qy 2577 taacctggaagaagacacagattatggaagctgctatgcaatcagcaagcttcatata 2636  
 Db 2812 CTCCTTTGAGAAAGACACACATCTTGGGGTCTCATTTTCTATACCCAGGCAATGATGTA 2871  
 Qy 2637 ttgtgctatgcagcaaggttgcattggaagcctatlaactaagctggaagcaagcaag 2696  
 Db 2872 TTTTTCATATGCTGCTGTTTCCGGTGGTGGTCTACTACTGAGGCAAAATGATGATGAA 2931  
 Qy 2697 cccagagggcatgtcatagtttcaatgcgaatgcatalgagagcatatgcagcaagcaag 2756  
 Db 2932 CTTTCAGAGATGTTCTTTGGTATTTCTCAGCATTAATGTTGGTGGTGCATGCGGAGCA 2991  
 Qy 2757 aacgctgcttgcgtcgtgaatatccaaagcaaatcgcgggctgcgcacatgcttgc 2816

Db 2992 GGTCACTTATTTGCTCTGACTATGCCAAAGCCAAAGATATCAGACGCCACATCATAT 3051  
 Qy 2817 ctgtgtgaaagaaacccaatatagacgcccagtcgaagaggaagaaagccagacac 2876  
 Db 3052 GATCAATTGAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCTTAACGCCAATATAC 3111  
 Qy 2877 atgtgaaggaatttagaagttcgaagagctccttctctcatalccatgctgcagagatg 2936  
 Db 3112 GTTGGAGAAATATGACATTTAATGAGTCTGTTTCAACTATCCACATCGACACAGACAT 3171  
 Qy 2937 ttcaatcctcgtggtcattccctcaagatgtgagcaggaagaaagcaagtaagatgtgag 2996  
 Db 3172 CCCCCTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAGGCCACAGCCTGGCCCTCTAGG 3231  
 Qy 2997 gaggcagcgctgtgggaaagaaagcaatcgttcaactctgcagagagacttatgacccgt 3056  
 Db 3232 TAGCATGCTGTGGAAGACAGATTTGACCTCTAGAGCGCTTTATGACCCCTT 3291  
 Qy 3057 gcaaggaacagctgctgtgagtggtgagtgcaaaagaaatgaaatgtaagtgctccg 3116  
 Db 3292 GGCTGTTCAATGCTATTTGATGAGGCAAAAGATTAAGACACCTGAATGTCCAGTGGCTCG 3351  
 Qy 3117 ttccaataagcaatcgtctcctcaagagcctgtgctcctcaactgcagcaatgctgagaa 3176  
 Db 3352 AGCACACCTGGGCATCTGCTCAGAGGCCATCTGTTGATGACGATGCGAGAA 3411  
 Qy 3177 catgctcatgtgtgacaagaagccgtgtgtgtgcatlaagtgagatcaagaagccgcaaa 3236  
 Db 3412 CATTCCTATGAGACACAAAGCGGCTGTATCATGTAAGAGATTGTTCAGAGCAGCAA 3471  
 Qy 3237 tgcgaataatccattcttattgaagtgctcccttgaagaataacacacaaagcttg 3296  
 Db 3472 GGAGGCCAACATACACACTTTCATGACGACACTCCTTGAGAAATACAAACACAGATAG 3531  
 Qy 3297 actgaaggaagcaagcttctcgtgcgcgcagaaacaaagactagatgtcaagggctct 3356  
 Db 3532 AGACAAAGGAACCCAGCTCTGTGTGTCAGAAACAGCCATTTGATGCTGCGCTCT 3591  
 Qy 3357 tctccaaaaccccaaatlttatgtgtgagatgagcgcaacttaagcctcgtataatgagc 3416  
 Db 3592 TGTAGACAGCTCATATTTTCTTTGATGAAGCTATCATGACTCTGATACACAAAAG 3651  
 Qy 3417 tgaagagtggtgtcgaagcctcgtgaagcctcgtgaagcaggaagagacatgagtg 3476  
 Db 3652 TGAAAAGGTTGTCCAGAAAGCCCTGACAAAGCCAGAAAGCCGACCTGCTTGTGAT 3711  
 Qy 3477 cactcaagcctctcgtcaatcagaacgcagaaattgtagtgtgtcgtcacaaatgaa 3536  
 Db 3712 GCGCCACGCTTGTCCACCATCAGAAATGCAATTTAATAGTGTGTTACAGATGCA 3771  
 Qy 3537 gataaaggaagaagaatcacaagagctcctgagaatcgagacatatatttaagtt 3596  
 Db 3772 AGTCAAGGAGCATGAGCACATCAACGCTGCGCCAGAAAGGATCTATTTTTCAT 3831  
 Qy 3597 agtgaatgcacagtcag 3613  
 Db 3832 GGTCAGTGTCCAGGCTG 3848

RESULT 4  
 AX105057  
 LOCUS AX105057 4279 bp DNA PAT 30-APR-2001  
 DEFINITION Sequence 1 from Patent WO0123540.  
 ACCESSION AX105057  
 VERSION AX105057.1 GI:13921209  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 dog.  
 Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 REFERENCE  
 1 (bases 1 to 4279)  
 Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and  
 Patten,C.J.



1719 tatataaaagcgtatga-acagatggaatcatatattactgaagaagacaa 1777  
1912 AAGAGAAATGAAATTTGAGTATGAAATGCCACTGTAATCAAAAGTGAATGC 1971  
1778 actcaatccctctgacactctgtgaagag-----catcaagtcagactcatgacaa 1829  
1972 CTTGGAATGCTCCAAAAGATTGAGGCTCCAGTTTAATAAAAGAAATCAACTCGCAG 2031  
1830 ggcggaagaaatcccaacaaatctaaagataagctct----- 1866  
2032 GAGTATACATGCACCAAGGCCAGACAGAAAGCTTGTAACAAAAGAGACTGAATGA 2091  
1867 -----cctgaagctctctataaataatttaagtttaacaagaagctgaagctctt 1919  
2092 GAATGTACTCTCCAGTTCTCTTGAGAGATTGTGAAGCTGAACCTGAACCTGAATGCGCTTA 2151  
1920 tctgtctctggaacatggtctctgtcttaaaatggaactgtctcaacagtaattccat 1979  
2152 TTTTGTGGTGTGATATTTTGTGCTATTTAATAAGAGAGCGCTCGAACCAGCATTTTCAT 2211  
1980 catctctgcaaaaatataaaccatgtt---ggaataatgataaaccacatlaagca 2036  
2212 AATATTTTCAAGGATTTATAGGATCTTACCAGATGAGATCCGTGAACAAAACGACA 2271  
2037 tgaigcagaatattatcatgatatctgcatcttgggtgtgtattgttgcagttta 2096  
2272 GAATAGTACATGTTTGTGATTTGTCTTACTGCTTGGAATTAATTTCTTTATTTACAT 2331  
2097 tctcaatgaggaatattctacgagagcaagcaagaaatttaagatbagaatlaagca 2156  
2332 TTTTCTCCAGGCTTCACATTTTGGCAAGCTGGGAGATCCCTACATACAGGCTTCGATA 2391  
2157 ctgtgctctcaagcaatgatatataatcagagataatgctgtgttgaatgaagaagaaacag 2216  
2392 CATGTTTTCAGATCCAGCTGAGACAGGATCTCAGCTGTTGATGACCTTAAACAC 2451  
2217 cacaaggaagctgacaaatattagcatalagatalagacaaatcacaaggaagcaagag 2276  
2452 CACTGAGACATTGACACACAGGCTTCCCAATGATCGGCTCAAGTTAAAGGGGCTATAGG 2511  
2277 ttcagaatctggtcgtcttaacacaaatgacaactacaatggaactctcaatcatctc 2336  
2512 TTCCAGGCTTGTGCTGATTAACCCAGAAATATAGCAATCTTGAGACAGCATTTATATAC 2571  
2337 ctctataatgagatggagatgacatctcctgaatcttgagatctgctcgaatctgagcgt 2396  
2572 CTTAATCTATGTTGGCAATTAACTTTTACTTACGCAATTTGAACCATATTTGCAAT 2631  
2397 gacaggaatgaltgaacacgacagcaatgactgattctgcaacaagaatlaagcaagact 2456  
2632 AGCAGGACTTGTGTAATGAAATGTTGTCTGACAGACAGCTGAAGAAAGATGAAGAGCT 2691  
2457 taagcaatctggaagaatagcaactgaagccttggagaataagctactatagtgatcatt 2516  
2692 AGAAGGACTGTGGAGATTGCTACAGAACCATCGAAGAACTTCGAACTGTTGTTCTTT 2751  
2517 aacaaggaagaaagccttcgagcaaatgatatgaagagatgcttcgaacacacagaa 2576  
2752 GACTCGGAGCAGAAATTGAATATCATATGACAGAGATTTTGCAAGTACATACAGAA 2811  
2577 taactcgaagaagaagcattatggaagcgtgtatgcatcagacagccttatata 2636  
2812 CTCCTTGAAGAAAGACACACATTTGCGGCTCATTTTCTATACACCGACCATGATGTA 2871  
2637 tcttgcataatgagcaaggttctgaattggaagcctataatcaactgagcgaatgac 2696  
2872 TTTTCTATGCTGCTGCTTTTCCGGTGTGCTACTACTGCTGAGCAAAATGAGTTCATGAA 2931  
2697 cccagaaggaatgctcatagtttctacgcaatgcatatggaagctatgacatgagaa 2756  
2932 CTTTCAGAGATGTTCTTTGATTTCTCAGACATATGTTGTTGCTGCATGGAGGCGCA 2991

2757 aacgcgtcttctgctcctgaatatctccaaagcacaatcgagggtctgcgcatctgttcg 2816  
2992 GGTGACTTATTTGCTCCTGACATATGCCAAAGCAAGTATACGACGCCACGCTATCAT 3051  
2817 ctgttgaagaagaacccaatatagacagcgcgacgtlcaagaagggaagaaagcagacac 2876  
3052 GATCATTTGAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACAGGCTCAAGCCAATATAC 3111  
2877 atgtgaaggaatttagatcttcgaaagctctctctcctatcactatgctccgaatgt 2936  
3112 GTTGGAGAAATGTGACATTTATGAGTCTGTTTCACTTATCCACTGACAGACAT 3171  
2937 tttcaatccctcgtggtcttccctcagtatctgagcgaagaaagacagtagcatcttg 2996  
3172 CCCCCTGCTCAGAGGGCTGAGCTCGAGGTGAAGAAGGCCAGACGCTGCGCTGATAG 3231  
2997 gacgacggtctgbyggaaaagcactctgttcaactctgcagagacttataagccgt 3056  
3232 TAGCAGTGGCTGGGAGAAAGCAGAGTTGTTGAGTCTAGAGCGCTTCATGACCCCTT 3291  
3057 gcaaggaacaaatgctgttggatgtgtgatatgaaagaatgaaatgataagtgatcagtgctcg 3116  
3292 GCTGCTTCAAGTCTAATTATGATGCAAAAGATTAAGCAACCTGAATGTCAGTGTCCG 3351  
3117 ttcaccaatagcaatcgtctcccaagagcctgtgctcctcaactgacagatctgagaa 3176  
3352 AGCACACCTGGGCTATGCTGTCTGAGAGCCCATCTGTTGACTGACGATTTGCCAGAA 3411  
3177 catgcctctatgtgaacaagcgtgtgtggtccatagatatgataaagaagccgcaaa 3236  
3412 CATTCCTATGAGACACAGCGGCTGATACATGAAGAGATTATGACAGCAGCAAA 3471  
3237 tgcagaaatatccatctcttctatgaaggtctcccttggaataatacaacaagtgtg 3296  
3472 GGAGGCCAACATACACACCTTCAATCGACACACTCCCTGAAATTAACACCGAGATGAG 3531  
3297 actgaagaagagcaagctctctggtcggaagaagaagacagctgcatctgaaggtctc 3356  
3532 AGCAAAAGGACACCACTCTCTGTGTGCGCAAGAACGCGCATTTGCTAGCTGCGCTCT 3591  
3357 tctccaaaaccacaaatattatctgttgaatgaagcgaactcagccctcgataatgacag 3416  
3592 TGTATACAGCCTCATATTTTGTGATGATGAAGCTACATCAGCTGATACAGAAAG 3651  
3417 tgaagaagtgtgttcagcatgaccttgataaagcagaaggaaggaagacatgcttagtgt 3476  
3652 TGAAGAGTTGTCCAAGAGCCCTGGACAAAGCCAGAGAGGCCGACCTGCAATTTGAT 3711  
3477 cactcaagagctctcgtcaaatcagaacgcaagattgtagtgtgtctgcacaaatggaag 3536  
3712 CGGCCACGCGCTTGTCCACCATCCGAATGCAAGTTTAATATGTGTGTTTCAGAAATGGCA 3771  
3537 gataaaggaacaaagaaatcacaagagctcctgagaatcgagacatalatttaagtc 3596  
3772 AGTCAAGGAGCATGCGCACATCAACAGCTGCTGGCCAGAAAGCATATTTTCCAT 3831  
3597 agtgaatgacagtcag 3613  
3832 GGTCAGTGTCCAGGCTG 3848

RESULT 5  
AF269224  
LOCUS AF269224 4045 bp DNA SYN 04-JUL-2000  
DEFINITION Synthetic construct Canis familiaris his-tagged-multidrug  
resistance glycoprotein gene, complete cds.  
ACCESSION AF269224  
VERSION AF269224.1 GI:8926216  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM artificial sequence.  
REFERENCE  
1 (bases 1 to 4045)



AUTHORS	Roullet, A.C., Puel, O., Gesta, S., Alvinerie, M. and Pineau, T.J
TITLE	Analysis of dog MOR1 p-glycoprotein
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 4045)
AUTHORS	Roullet, A.C., Puel, O., Gesta, S., Alvinerie, M. and Pineau, T.J
TITLE	Direct Submission
JOURNAL	Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de Tournefeuille, BP3, Toulouse 31931, France
FEATURES	Location/Qualifiers

```
source      1..4045
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="original mRNA isolated from Beagle dog colon"
DSD         101..3964
```

```

/codon_start=1
/transl_table=11
/product="his-tagged-multidrug resistance glycoprotein
MDR1"

```

BASE COUNT	1171 a	850 c	986 g	1038 t
ORIGIN				

Query Match	36.58;	Score 1332.2;	DB 12;	Length 4045;
Best Local Similarity	62.68;	Pred. No. 1.7e-267;		
Matches 2188; Conservative	0;	Mismatches 1258;	Indels 51;	Gaps 6;

Oy	162	cagtgtaacctgtatatagtttggaatagtggtgtgccttgatttttgytacaaca	221
Db	460	CACGATGCCCTATTATTAACAATGGGATCGGGCTGGGCTGGTGGCTCTTACATCCA	519
Oy	222	gatttcctgtggatttaactcagcagcaacacccaagagatttcgaacaacgtttc	281
Db	520	GGTTTCATTCTGGTCCCTGGGACGAGAACACATATCTCAAAATTGAAAACATTTTT	579
Oy	282	tcaatcagttttggcaacagacatcggctgtgttgaatagctgtgacatcggtaactaa	341
Db	580	TCATGCTATCATGGACAGAGATGTTGGCTGTTTGACGTGCATACGTTGGGGAGCTTAA	639
Oy	342	cactgcagatgaacagacatg---acaatatcagatgagatctggaataaagtgtctc	398
Db	640	CACCGGCTCACAGCGATGCTCCAAATATAAGAGGAATTTGGCAAAAATTTGGAAT	699
Oy	399	gttgttcaaaaacatgctactcttttcgatttggcgtcgacgttggttggltgaagagctg	458
Db	700	GTTCTTTCAATCATAGACAAATTTTACCGGTTTATTAAGTGGGGTTTACACGTGGTTC	755
Oy	459	gaactcacccatgagctctatcaacagttctcctcttataatggtcttcacgcgcagcatg	518

Db	760	GAAGCTAAACCCCTTGATGATTTTGGCCATCATGACCCCTGTCTTGAGACTTTGACGGCCATCG	819
Oy	519	tcttagtgatgcatcctcaatgaccagaatgaatlaagtgctcatlccaaagctggggc	578
Db	820	GGCAAGATACATCTCTTCAATTTAGATATAAAGACCTTGCGCTATGCAAAAGCTGGAGC	879
Oy	579	tgtggcagaagaagctctgtgcatcaatccgaacagtcataagctttagggcccgagaga	638
Db	880	AGTACTGTGAAGAGCTTGTAGCAGCATCGAAGCTGTGATTGGCTTTGGAGCAAAAGAA	939
Oy	639	agaacttcaaagtgtaaccgagaacttcaaagatgcaagagatttggcataaaagagac	698
Db	940	AAACCTGTGAAGGTACACAAAATTTAGACAGAGCTMAAAGATTTGGATTAAGAAAC	999
Oy	699	tatagcttcaaagaagtgcctctgtgtgctgtgactcttatalgaaatgaaactatgagct	758
Db	1000	TATCAAGGCCAACATTTCTATTGGTGCCCTTCTTATGATCTATGCAATATGCTCT	1059
Oy	759	tgctcttggatlgaaacctccttgatcttcaatgagaacctggaatataccatcgagac	818
Db	1060	GGCTTCTGTATGGAGACCTCTGTGGCTCTCCAA----TGAATATTTCTATTGGACA	1113
Oy	819	tgctctgtgtttcttctttagtgatataccatagcagtatatgcatggagcaagctcc	878
Db	1114	AGTACTCATGCTCTTCTTTCTGTATTAATGGGCTTTAGATTTGACAGGCAATCCCC	1173
Oy	879	tcaacttgaacacctgcgcataagcccgagagctgtccttcatalcttccagttatga	938
Db	1174	AAGCATTTGAAGCATTTTGCAAAACGACAGAGAGCAGCTTATGAATTTCCMAAGATATTTGA	1233
Oy	939	taagaaaccagtatatgaataccttccacagctgatataaacctgaatccatagaaag	998
Db	1234	CATAAACCAGCATTTGACAGCTATTGGAAGAGTGAACATMAACAGATATATTTAAGG	1293
Oy	999	aactgtggaatttaaaatgatttcttccataatccatcaagacctcatcaagattct	1058
Db	1294	AAATTTGGATTTAAAAATTTTACTTTCAGTTACCTTCTCGAAAGAAGATTTAAGATCTT	1353
Oy	1059	gaaaggctgaactcctagaatlaagctgtggaagacagctgcctgtgctgcgtcctaag	1118
Db	1354	AAGGGCTCAACCTGAGAGTTCCAGATGGGCGACAGAGTGGCCCTGGTTGGAAACAGTGG	1413
Oy	1119	cagtggaagaagtaagctagtcacagctctgcagaggtatatagcccgatgaigagct	1178
Db	1414	CTCGGGGAAGACGACGCCCTGTGAGCTGATGACAGAGGCTCTATGACCCCAAGATGGCAT	1473
Oy	1179	tatcatagtgtagtagaagaatgacatcaagagctttaaagctgcygcatatcatagacaat	1238
Db	1474	GGTCTGATGATGAGCAGACACATTTAGGACCATTAATGTAAGCATCTTTGGGAAATTAC	1533
Oy	1239	tggaatgtagttagtcaagaagccttltgtctggagaccacatcaagtaacaatatcaagta	1298
Db	1534	TGATGTGGTAGTCAGAGAGCCTGTTGTTCACACATATAGCTGMAAAACATTTCCCTA	1593
Oy	1299	tggaacgagatgtagtgaactatgaagagatgagagagagcagcaagggaaacaaatggtga	1358
Db	1594	TGGCCGGGAAAGTCCACATGGATGTGAGTTGAAGAAAGCTTTAAGGAAGCCAAATGGCTA	1653
Oy	1359	tgatttatcatatgagtttccctaataatlaatactatgtagttaggggaaaaagagtgca	1418
Db	1654	TGATTTTATATGATAACTACTTAATTAATTTGACACCTGTGGTGGAGAGAGAGGGCCCA	1713
Oy	1419	aatgagtggaagggcagaagaacagagatgcgaattgtcgtgcttagtctgaaaccccaa	1478
Db	1714	GGTAGAGTGGAGACAGAAACAGAGAAATCGCATTTGCTGGGGCCCTGTGGTCCAAACCCCA	1773
Oy	1479	gattctgatttagatgaagctacgtctgcctctgattctgaaagcaagtcagctagtcca	1538
Db	1774	GATCTCTCTGCTGATGAGGCAAGCTCAGCTCTGGAACCTGAAAGTGAAGAGAGTGGTTCA	1833
Oy	1539	agctgcaactgagaagcaggaacaaagtcgagatacaaatgttgtagcaacagctatcc	1598
Db	1834	GGTGGCCCTGGATTAAGGCCAGAAAAGCGGAGCTACCAATTGTGATAGCTCATTCGTTTCTC	1893



Oy	1599	tactatcgaagtgcaagattgatgtgacccctaaaggaatgctgcgcggaagaag	1658	Oy	2637	tttgccatgcagcagggttcgaattggaagccatttaacttaactgaagcagatgac	2696
Db	1894	TACAGTTCTGTAATGCCAGTGCATTTGCTGTTTGATGATGAGATGATGGAGAAAGG	1953	Db	2974	TTTTTCCATATGCTGGCTCTTTCCGGTTTGGTGGCCACTTGGTGGCAATGATTCATGAA	3033
Oy	1659	agccatgctgaactaaatgycanaaacagagtcataatattcactctgtagtcaaga	1718	Oy	2697	cccgaaggaatgctatgactatgatttactgcaattgcatatgagagcagcagcag	2756
Db	1954	AAATCATGATGACATCATGAAAGAGAGGATTTACTTCAAACTGTGCACAAATGCACAC	2013	Db	3034	CTTTCAGGATGTTCTTTTGGATTCTCAGTATTTGTCTTGGTGGCCATGCGCAGTGGGCA	3093
Oy	1719	tattaaagaagcgtatga-acagatggaagtcataatgacatattctactgaaagaacca	1777	Oy	2757	aagcctgcttgctccttgatattccaaagcccaacccgggggtgctgcatctgtgc	2816
Db	2014	AAGAGGAATGAAATTTGATTAGAAATGCCACTGTGAATTCAAAAGTGAAGGATGC	2073	Db	3094	GCTCAGTTCAATTTGCTCTGATGATGCAAAAGCCAAAGTATCAGACGCCACGTATCAT	3153
Oy	1778	actaactcctcgtgacccgtggaag-ccatcaagtcgaagctcatgagaca	1829	Oy	2817	cttgctggaagaagaacaaatataagacagccgcaagtcgaagaaggaagccagac	2876
Db	2074	CTTGGAAATGCTCTCAAAAGATTCAAGGTCAGTTTAATTAAGAAAGATCACTGCAG	2133	Db	3154	GATCTTGAATAAAAGCCCTTGATTTGACAGCTACAGCCCTCAGGCCCTCAAGCCAAATAC	3213
Oy	1830	gactgaggaatcccaactaaagataagtcct-ccatcaagtcgaagctcatgagaca	1866	Oy	2877	atgtgaaggaattagaagtttcgagaagctcttcttctatacgaatgcgccagatgt	2936
Db	2134	GAGTATACATGCACACAGGCCAAAGACAGAAAGCTTGTCAAAAGAGGACTTGAATGA	2193	Db	3214	GTTGGAGGAATGATGACATTTAATGAGTGTGTTCAACTATCCACTGCACAGACAT	3273
Oy	1867	-----cctgaagtcctcattataaaatlttaaglttaacaagcctgaatgcctt	1919	Oy	2937	tttcaactcgttggtctatccctgaatgagcgaaggaagaagatgacatgtgtgg	2996
Db	2194	GAATGTACCTCCAGTTTCTCTGAGAGATTCTGAAAGCTGAACCTCAACTGATGGCTTA	2253	Db	3274	CCCTGCTCCAGGGGCTGAGCCTTGAGGTGAAGAGGCCAGAGCGTGGCTGTAGG	3333
Oy	1920	tgigtcttgaggacattgcttctgtcttaaggaactgttcatcagatattccat	1979	Oy	2997	gagcagcgctggtggaagaagcaactctgttcaactctgcgaggaacttatgaccgt	3056
Db	2254	TTTTGTGTTGTAATTTTGTCTATTATTAACGAGGCTTGCAACGACGATTTTTCAT	2313	Db	3334	TACGATGCTGTGGGAAGACACAGTTGTTACGTTCTTAGACGCTTCTATGACCCCTT	3393
Oy	1980	catcttgcnaaaatlataacacatgtt---gaaataatgataaaacacatlaaaca	2036	Oy	3057	gcaaggaacatgtcgttctgtagtgcgtgtagtgcnaaaagaatgaaatgacatgctgcg	3116
Db	2314	AATATTTTCAAGATTAATAGGATCTTACCAGATGAGATCCTGAACAAAGCA	2373	Db	3394	GGCTGCTCAGTGCTAATTTGATGCGCAAGATTAAGACACCTGATGTCAGTGGCTCG	3453
Oy	2037	tgatgcagaatlttaaccatgatactgcatacttgggtgttattctgtctgtcagta	2096	Oy	3117	ttcccaatagaacatgcttctcctcaagaagcgtgtgtcttcaactcgaactgtgcga	3176
Db	2374	GAATAGTAACTGTTTCTGATTGTTTCTGATGCTTGGAATTAATTTTATTATACAT	2433	Db	3454	AGCACCTCGGACATCGTCTCAGAGCCCATCTGTTTGATGCTGCACATTTCCGGAAT	3513
Oy	2097	ttcaatgaggaattatttccgcagagcaggggaatlttaagaatgaaatgaaca	2156	Oy	3177	catgcctatgtgacaaacagccgtgtgtgcattagatgataagacaaagaccgca	3236
Db	2434	TTTTCTCCAGGGCTTACATTTTGGCAAGCTGGGAGATCCTCACTAAGCGCTTCGATA	2493	Db	3514	CATTGCCATGAGACCAACAGCGGGTGTGTATCATATGAAGATTAATGACAGGCCAA	3573
Oy	2157	cttgacctcaaaagcatgtatatacagataatgcgtgttgaatgaagaagaacag	2216	Oy	3237	tgacgaaatatacattcttattgaaggtcctccctgagaatacaacaacaagtgtg	3296
Db	2494	CATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGTGTGTTGATGACCCATAAAGAC	2553	Db	3574	GGAGGCCAATATACACCACTTCATGAGACACTCCTGAGAAATFCAACACAGAGTAGG	3633
Oy	2217	cacagaggtcttgacacaataatagccatagatagacaaatltcaaggaagcaag	2276	Oy	3297	actgaagagacagcgttcttgccgcgcgaagaacaagaacagctatgtcaagcctc	3356
Db	2554	CACGTGACATTTGACAAACAGGCTTGGCAATGATCGGCTCAAGTTAAAGGGCTATAGG	2613	Db	3634	AGACAAAGAACCCAGCTCTGTGTGGCCAGAAACAGGCAATTCATAGCTCGGCTCT	3693
Oy	2277	ttcgaaggttgcttcttaacaacaatgcacactaacatggaacttcaagtatattc	2336	Oy	3357	tctccaanaacccaatlttatgtgtgtagagccacttgaagcctcgaataagacag	3416
Db	2614	TTCCAGGCTTCTGTCATATCCAGAAATATAGCAAAATCTTGGGACAGCCATTTATATTC	2673	Db	3694	TGTTAGACAGCTCATATTTTGTGTTTGGATGAACTTACATCAGCTTGGATACGAAAG	3753
Oy	2337	cttataatagatgagatgacatctcgtatcttgatctgtcagtaactgtgcgt	2396	Oy	3417	tgaagaagtggttccagcatgccccttgataaagccagacggaagaagatgctagtg	3476
Db	2674	CTTATCTATGTTGGCAATTAACACTTTTACTCTTACCAATTTGACCATCATTCGCAAT	2733	Db	3754	TGAAGAAGTTTGTCCAAAGACCTCGACAAAGCCAGAGAGAGGCCACCTGATGTGAT	3813
Oy	2397	gacaggaatgtgaacccgacgaatgactgattgccaacaagaataagaaact	2456	Oy	3477	cactcaagagctctctgcaattcagaacgcagattgtatagtggtctgcacaatgaaa	3536
Db	2734	AGCAGGACTTGTGAAATGAAATGTTGTCTGACAAACACTGAAAGAAAGTAAGAAAGACT	2793	Db	3814	CGCCACCGCTTGTCCACCATTCAGAAATGCAATTAATAGTGGTTCACAAAGGCA	3873
Oy	2457	taagcatgctggaagaatagaaactgaaagccttggagaatatagttactatagtgctt	2516	Oy	3537	gataaaggaacaaggaatcactcaagagctccttgagaatcgaagacataatttaagt	3596
Db	2794	AGAAAGACTGTGGAAGATGTTGATACAGAACATCGAAAACCTTCCAACTGTGTTCTTT	2853	Db	3874	AGTCAAGGAGCATGCGACACATCAACAGCTGCTGAGTCAAGAAAGCATTTTTCAT	3933
Oy	2517	aacaaggaagaagccttcgagcaaatgtatgaagagaatgcttcaagactcaacagaa	2576	Oy	3597	agtgatgcacagtcag 3613	
Db	2854	GACTCGGAGAGCAAGTTTGAATACATGTATGACAGAGTTTGGCAAGTACCATACAGAAA	2913	Db	3934	GATCAGTGTCCAGGCTG 3950	
Oy	2577	tactcgaagaagcagatattggaagctgtatgcatcagcaatgccttatata	2636				
Db	2914	CTCTTTAGGAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTA	2973				

RESULT 6  
AF045016 AF045016 4317 bp mRNA MAN 07-FEB-1998  
LOCUS Canis familiaris multidrug resistance p-glycoprotein (MDR1) mRNA,  
DEFINITION



Qy	1479	gattcttgatctttttagatgagagctcgtctcgtcccttgattccagaagaagcaagtcagctgttca	1538
Db	1722	gattcttcttctctctggatgagagcaaacgctcacctctggcacctgaaatctgaagcagctgtttca	1781
Qy	1539	agctgcacacttgagaagaagcgagacaaagctggagactcaacatcggtgagacacgacttcc	1598
Db	1782	ggtagccctctggatgaagcccgaaaaagccggactccatcttgatgtagctcattcgtttgtc	1841
Qy	1599	tactattccagaagtgcagatttgaattgcacctaaagatggaatgcgtcggcgagaaag	1658
Db	1842	tacagcttcgtaatgccgagatgtctattctcgttttgatgatgagatctgttgagaaaaag	1901
Qy	1659	agcacatgctgaaactaatgycacaaacgagtgctatataatttcaactgtgatgcacaga	1718
Db	1902	aaatcatgatgaactcatgatgaagagaaagggcattttacttcaaacctgttcacaaatgcagac	1961
Qy	1719	tattaaaaaagctgatgga-acagatgtagtcaatgatcatattctctctgaagaagaagcca	1777
Db	1962	aagagcaaaatgaattgagttgaataaaatccactgctgatcccaaaatgtgaagtgatgc	2021
Qy	1778	actaactctctgcgaactcgttgaagag-----catcaagtcagaactcatatgacaa	1829
Db	2022	cttggaatgtctctccaaaaagatttcagagctccacttttaataaaaaaagatcaactcgcag	2081
Qy	1830	ggctgtaggaatccaccacctaagaagata-----	1860
Db	2082	gagatattcatgacacccaaagggccaaagcaaaagccttgctacaaaagagacttgaaatga	2141
Qy	1861	-agcttctccgaagtcctctctatataaaattttaaagttaaacaagccttgaaatgacctt	1919
Db	2142	gaattgtctctcaggttcccttctggagagattctgaagctcaaacctcaactgaaatggcctta	2201
Qy	1920	tgtgtctctgggagacattgacctctgtcttaataatggaactgtgtctccagatatttccat	1979
Db	2202	tttttgctgttgatattttttgtgcatttttaaacagagagccctgcgaacacagactttttcat	2261
Qy	1980	catctctgcacaaactatataccaagttt---ggaaataatgaataaacacatataaagca	2036
Db	2262	aatatttttcaagatttatatagggattcttttcccgagatgcagattcctgaacaaaaacga	2321
Qy	2037	tgatgcagaatttatccatgatgatctgtcaatttgggtgtatttgccttgccttgcagtta	2096
Db	2322	gaatagtaaacattgtttctgtattgtttgttaactccttggaattatttttctttttattacatt	2381
Qy	2097	tttcatgcagggatttttaagcgcagcgacagggaaattttaaagctgagatataagca	2156
Db	2382	ttttccttcacagggcttcacatttttgccaaacctggggagatcctcacacttaacggctttcgata	2441
Qy	2157	cttgaccttcaaaagcaatgtlatatcagaatactgcctgtgttgaatgaagaagaacaag	2216
Db	2442	catgctgttccagatccatgcctgcagacagagatgcagctggtttgattgacaccttaaaaacac	2501
Qy	2217	cacagggagctgtgaacacaaatattagccatagatatagcacaaattcaagaagcaacagg	2276
Db	2502	cactggagacatttgacaaacacagccttccaatgatgcgctcaagttaaaagggcgctatagc	2561
Qy	2277	ttccagagattgggctcttaacacaaatgcaactcaacatcagcaggaacttcagttatcattc	2336
Db	2562	tttcagagcttgcgtctattttaccacgaattatagcaaatcttggagacacagcattttattatgc	2621
Qy	2337	cttataatagatgtaggagatgacatccctgtgattctgagttatgtctcagttactgtccgt	2366
Db	2622	cttttaattatggttgccaaatttaaacactttttactttagcaattgtttaccatcatattgcatt	2681
Qy	2337	gacaggaatgatttgaaaacgcagcaatgtgacttgatttgcacaaagaataagcaagaact	2456
Db	2682	agcagggaggttcttgaatgaataatgtttgcttgacacagcactgaaagaatgaagaaagact	2741
Qy	2457	taagcatgtctggaagaagtagcaactgaagcttcttgagaaataatgactatagttgactt	2516
Db	2742	agaaaggagcttgggaagattcttctacaaagccatcgaaaaactttccaaactgtgtgtttcttt	2801
Qy	2517	aacaaaggaaanaagccttcgagcaaaatgtlatgaagaagatgcttccagactcaacacaga	2576

Db	2802	GACTCGGAGCAGAAAGTTTAAATACATGTAATGACACGAGATTGGCAAGTACCATACGAA	2861
Oy	2577	tactcgaagaagaacacagaattatctbgaagccgtctatgcattccagccattcata	2636
Db	2862	CTCTTTTAGGAAAGACACACATCTTCGGGGCTCATTTTTCATACCCAGCAATGATGTA	2921
Oy	2637	ttctgcctatgcagcagaggttcgcattctggagcctatctaactcaagctgcagatgac	2696
Db	2922	TTTTTCTATAGCTGGCGTTTCGGGTTTGGTGCTCACTTGGTGGCAAAATGACTCATGAA	2981
Oy	2697	cccagaggcaltgcatagtttcttctacgcgaattgcatactggactatgcccactgaa	2756
Db	2982	CTTTTCAGGAGTCTCTTTTGGTATTTTCAGCTATTTGCTTTGGTCCATGGCAGTGGGCA	3041
Oy	2757	aagctcgttttgcctctbcaatctccaagcgaacatcgsgggtctgcacatctgttc	2816
Db	3042	GGTCAGTTCAATTTGCTCTCTACATATCCAAAGCCAAAGATACAGACGCCATCATCAT	3101
Oy	2817	ctctctggaagaagaaccaaatatagacagccgacatcaagaaggaaaaagccagac	2876
Db	3102	GATCATTTGAAAAAGCCCTCTGATTTGACAGCTACACCCCTCACGGCCTCAAGCAATAC	3161
Oy	2877	atgtgaagggaatttagagttctgagaagctctcttctctatcgaatgcgccagatgt	2936
Db	3162	GTTTGAAGGAATGTGACATTTAATAGGTCGTGTCGAACATCTCCACATCGCACAGACAT	3221
Oy	2937	tttatcctctcgtgcgtactccctcagatcttgagcgagaagaagaagtagcattgtgg	2996
Db	3222	CCCCGTCTCCAGGGGCTGAGCCTTCAGAGTGAAGAAAGGCCACAGCCTGGCTGTAGG	3281
Oy	2997	gagcagcgcgctgtggaagaagcactctctcaactctccagagagacttaatgacccgt	3056
Db	3282	TAGCAGTGGCGTGGGAAGACACAGTTGTCAGTCTTAAGAGCGGTTCTATGACCCCTT	3341
Oy	3057	gcaagagcaagctgcgtcttgaatgtgtgatactgcaaaagaattgaatgtaacgtgtccg	3116
Db	3342	GGCTGTTTCAGTGTATTTGATTTGGAAGAAAGATAAAGCACATGAAATGTCACATGCTCCG	3401
Oy	3117	ttcccaaatgcaatctgtctccataagaagcctgagctctcaactcagatctgcgtgaa	3176
Db	3402	AGCACACCTGGGCACTGTGCTCTGAGAGCCCATCTCTTTGACTGCGAGATTGGCGA	3461
Oy	3177	catgcctatgtgacacaagccgctgtgtgcattagaatgataccaagaagccgcaaa	3236
Db	3462	CATTGCTATGAGACAAACAGCCGGGTGTAATCAATGAAAGATTTATCCAGCAGCCAA	3521
Oy	3237	tgacgcaaatatccatctttatctgaagctctcccttgaaatacaacaacaagaattg	3296
Db	3522	GGAGGCCAACATACACCACTTCATTCGAGACACTCCCTGGAATAATCAACACACAGATGAG	3581
Oy	3297	actgaagaaggagcaagctcttcctgcggccagaacaagaacacagctatctgaaggctc	3356
Db	3582	AGACAAAGGAACCCAGCTCTCTGTGTGTGGCCGAAGACGCCATATGCATAGCTCGGGCT	3641
Oy	3357	tctccaaaaccocaaatttatactgttgaatggagccacttaagccctcgataatgac	3416
Db	3642	TGTTTAGACAGCTCATTTTGTCTTTTGATGTAAGCTACATCAGCTCTGATACGAAG	3701
Oy	3417	tgagaagtggtctcagacatgccttgataaagccagagcaggaaagacatgcttagtgt	3476
Db	3702	TGAAAAGGTGTCCAAAGAACCTCTGACAAAGCCAAAGCCAGAGGCCGACCTGCATGTGAT	3761
Oy	3477	cactcacaagctctctgcgaattcagaacgagaattgtatagtggttctgcaaaatggaa	3536
Db	3762	CGCCACCGCTGTGCCACCATCCAGAAATGAGATTTAATGATGTGCTTCACAAATGGCA	3821
Oy	3537	gataaagaaacaagaactcatcaagctccctgagaatactgagacataatcttaagt	3596
Db	3822	AGTCAGAGAGATGGCACATATACAGCTGCTGGCTCAGAAAGGCAATATTTTTCAT	3881
Oy	3597	agtgatgcacagtcag 3613	



Dh 1662 GCTGATGCTGGACAGAAACAGAGAAATGCCATTCCTCGGGCCCTGTGTTCCGACCCCA 1721  
Oy 1479 gatttcgttttagatgtagtgctacgtctgccttgatltccaagaacgaagtcgcttca 1538  
Dh 1722 GATTTCCTGCTGGATGAGGCAACGTCAGCTTGACACTGAAATGGAAGCAGTGGTTCA 1781  
Oy 1539 agctcactgtagaagcgagcaaaagtcgaactcaactcgtggtgagcaacacgacttc 1598  
Dh 1782 GGTGGCCCTGGATTAAGGCCAGAAAAGCCGGACTACCATTTGATAGCTCATGCTTTGTC 1841  
Oy 1599 tactatcgaagtcgaagtcgtatgtgtgtgacccctaaaggaatggaatgctggtcgagaaag 1658  
Dh 1842 TACAGTTGCTGATAGCCGATTCATTTGCTGTGTTGATGATGAGATCATTTGTGCAAAAG 1901  
Oy 1659 agcaatgctgaactcaatggtgcaaaacgaagtcctatattatcactctgtgtgacgaaga 1718  
Dh 1902 AAATCATATGAAACATGAAAGAGAGAGGCGCATTTACTTCAAACTGTCCACATGCAGAC 1961  
Oy 1719 taataaaaagctgtatga-acagatggaatcaatgacatattcctaagaagaagacca 1777  
Dh 1962 AAGAGGAATGAAATGAGTTAGTAAGAAATGCCACTGTGATTCAAAAGTGAAGTGTATGC 2021  
Oy 1778 actcactcctcctgcactctgtgaaagag-----catcaagtcgaactcattgacaa 1829  
Dh 2022 CTTGGAATGCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAGAGATCAACTCCGACG 2081  
Oy 1830 ggtcgaagaaatccaccacatctaagaagata----- 1860  
Dh 2082 GAGTTACATGACACCAAGGCCAAGACAGAAAGCTTGTAACAAAAGAGACTTGATGA 2141  
Oy 1861 -agcttcctgaagtcctcctctctaaataatttaaagtaacaagcgtgaatgctt 1919  
Dh 2142 GAAATTCCTTCAGTTCTCTCTTGAGGATTTGTGAAGCTGAATCTGAATGAGCTCTTA 2201  
Oy 1930 tgtgttctggtggaacatggtctctgtcttaaatggaactgtctacagatattccat 1979  
Dh 2202 TTTTGTGTTGTTATATTTTGTGATATATAAGGAGGCGTGCACAGCATTTTTCAT 2261  
Oy 1980 catcttgcataaataatacaacatgtt---ggaataatgataaaacacataaagca 2036  
Dh 2262 AATATTTTCAAGGATTAATAGGATTTAACCCAGATGAGATCTGTGAACAAAAGCACA 2321  
Oy 2037 tgatgcagaataatttcaatgatatctgcatcttggtgtgtatctgtctgtcaatga 2096  
Dh 2322 GAATGTAACATGTTTCTGTAATTTTCTAGTCTTGGAATTAATTTCTTTATTTACAT 2381  
Oy 2097 ttcaatgcaggaattatatttcaagcagagcaagggaaattttaaagatbagaatgaaca 2156  
Dh 2382 TTTCTCCAGGGCTTCACATTTGGCAAGCTGGGAGATCCTCATAAGCGGCTTCGATA 2441  
Oy 2157 ctgtgcttcaaaagcaatgatatatacaggaatatgcctgtgtttgattgaagaagaacaag 2216  
Dh 2442 CATGCTTTTCAATCCATGCTGAGACAGAGTCTGAGCTGTTGATGACCTTAAACAC 2501  
Oy 2217 caccagagagctgacaaataatattagcaatagatagcaaatcaatcaagggcaacag 2276  
Dh 2502 CACTGGAGCATTTGACACACAGGCTTGCCANTATGCGGCTCAAGTTAAAGGGCTATATAG 2561  
Oy 2277 ttccaggaatggtcgtcttaacacaaatgcaactaacatgagacttcaatcatctc 2336  
Dh 2562 TTTCCAGGGCTGCTGTCATTTACCCAGAAATATAGCAATCTTGGACAGGATTTATATATC 2621  
Oy 2337 ctctatataatgtagtgaggaatgacatctcgtatctcgtatgtctcgaagtgcgt 2396  
Dh 2622 CTTAATCATATGCTTGGCAATTAACATTTTACTTTAGCAATTTGTAACCATATTCGCAAT 2681  
Oy 2397 gaccagaaatgtagaaacccgacgaactgactgattggtccaaacaaagaagaagaact 2456  
Dh 2682 AGCAGAGTTGTTGAAATGAAATATGTTGCTGACAGAGACTGAAAGATTAAGAAAGAGT 2741  
Oy 2457 taagaatgctggaagaatagcaactgaagcttggagaatatacgtactatgtcatt 2516  
Dh 2742 AGAAGAGCTGGGAAGATTTGCTACAGAAAGCCATCGAAAACCTTCGAACTGTTTCTTT 2801

---

Oy 2517 aacaagggaanaagccttcgagcaaatgatatgaagaatgcttcgaactcaacaaagaa 2576  
Dh 2802 GACTCGGAGCGAAGATTGTAATACATGATACAGAGTTTGCAAGTACCTATACAGAA 2861  
Oy 2577 taaccgagaagaagcacagatatttgaagatcgttatgcatcagcactgactatata 2636  
Dh 2862 CTCCTTTGAGGAAGACACATCTTGCGGTCATTTTCTATACCCAGGCAATATATGA 2921  
Oy 2637 ttgtcctaagcagaaggttctgatttgaagcctatttaaltcaagctgagcaatgac 2696  
Dh 2922 TTTTTCATATCGCTGCTTTTCCGTTTGTGCTCTACTTGGTGGCAAAATGACTCATGA 2981  
Oy 2697 cccaagggcaatgcatagatttttactgcaatattgacatataggaatgacatcggaa 2756  
Dh 2982 CTTCAGATGTTCTTTTGTGATTTCTCAGCTTAATTTCTTTGGTGCATGCGAGTGGCA 3041  
Oy 2757 aacgtctgtttgtcctcgaatatltccaagaacaaatcoggggtcgcacatctgtgc 2816  
Dh 3042 GGTCAATTCATTTGCTCCTGACTATGCCAAGCCAAAGTATCAGCAGCCACGTCATCAT 3101  
Oy 2817 ctgttggaaaagaaacaaatataagcaagccgagtcgaagaaggaaagccagacac 2876  
Dh 3102 GATCATTTGAAAAAAGCCCTGATGACAGTACAGCCCTCAGGCTTCACAAAGTAC 3161  
Oy 2877 atgtgaagggaattagagtttcgagaagctccttctctatccatgtcgcacagatgt 2936  
Dh 3162 GTTGAAGGAATGTGACATTTAATGAGTGTGTTCAATTTCCACTCCACGACGACAT 3221  
Oy 2937 ttcatcctcgtgtgcttatccctcagtatgtgagcgaagaagaagacagatgactgtggtg 2996  
Dh 3222 CCCGCTGCCAGGGGCTGAGCTTCGAGTGAAGAAAGGCCAGAGCTGCGCTCTGTAG 3281  
Oy 2997 gaagcagcgtctgtggaagaaagcactctgttcaactctcgcagaacttataagaccgt 3056  
Dh 3282 TAGCAGTGGCTGTGGGAAGAGACAGATTGTTCAAGCTCTAGAGCCCTGTATAGACCCCT 3341  
Oy 3057 gcaaggaacagtgtcgtgtgtatgtgtgtgatagcgaagaagaatggaatgacatgctcgt 3116  
Dh 3342 GCGCTGTTCCAGTGTCAATTAATGATGGCAAAAGATTAAGACCTGAATTTGCCATGCTCG 3401  
Oy 3117 ttcccaaatagcaatcgttccccaagaagcctgtgtcttcaactcagcatgacatgctgaga 3176  
Dh 3402 AGCACACTGGGCATGCTGTCTCAAGAGCCCATCTGTTTGTCTGACGATTCGCGAGAA 3461  
Oy 3177 catcgcctatgttgaacaaagccgtgtgtgcatatagatgaaacaaagaaagccgcaaa 3236  
Dh 3462 CATTCGCTATGAGAACAAACAGCCGGGTCTATCACATGAAGATTAATGACAGGCCAA 3521  
Oy 3237 tgcagcaaatatccacttcttattgaaggtctccctgagaataatcaacacaaagttgtg 3296  
Dh 3522 GGAGGCCAAACATACACACACTTTCATGAGACACTCCCTGGAATATCAACACAGAGTAG 3581  
Oy 3297 actgagaagagcagagcttctgcggtccagaagaacaaagactcagctatgcaaggtctc 3356  
Dh 3582 AGACAAAGGAACCCAGCTCTCTGTGTCGCGAGAAACAGGCGCATTTGCCATAGCTCGGCTCT 3641  
Oy 3357 tctccaaaacccaataatttatttgttgaatgaggccacttcaagcctcgaataatgacag 3416  
Dh 3642 TGTATGACAGCTCATATTTTGTCTTTGGATGAAGCTATACATCAGCTCTGATACAGAAAG 3701  
Oy 3417 tgaagaagtgtgtcagcatgcccctgtgataaagccaggaagggaagacatgctcgtgtgt 3476  
Dh 3702 TGAATAAGTTGTCCAAAGAGCCCTGGACAAACAGAGAGCCGACCTGATTTGTAT 3761  
Oy 3477 cactcaagagctcctgtgcaatcagaacgcagatctgtagagtggtctgtgcacaatgga 3536  
Dh 3762 GCGCCACGCTTGTCTCACCATTCACAAATGCAGATTTTAATAGTGTGTTTCAAAAGGCA 3821  
Oy 3537 gataaaggaaacaaagaaactcaagagctcctggaagaatcagagacataatattaaagt 3596  
Dh 3822 ACTCAAGAGCATGCGACACATCAACAGCTGCTGCTCGAAGAAAGCATCTATTTTTCAT 3881





QY	1860	aagccttcacggaagtcctctcctatataaaattttaaagtaaaacaagctcgatgacctt	1913
Db	2494	AAGATATACCTCCAGCTTTCCTTTGGAGATATATGAAGCTAAATTTTAACGTAAATGGCTCTTA	2553
QY	1920	tgtgtctctgggacattggtctctctcttaaatggaactgtctacccagatttccat	1979
Db	2554	TTTTGTGTGGTGTATTTTGTCTCATTATTAATGAGGCCCTGCACACGACATTTTGCAT	2613
QY	1980	catccttgcataaataataaccatgltt--ggaataatgataaaccacataanaagca	2036
Db	2614	AATATTTTCCAAAGATATATAGGGGTTTTTCAAGAAATGTGATCTCTGAACAAAACGACA	2673
QY	2037	tgtatgcgaattatttccatgatattgtcatttgggtgtattatgtcttctgtcagtta	2096
Db	2674	GAATATGTAACCTGTATTTTCACTATTGTTTGTACCCCTGGAAATATTTTCTTTATATCAT	2733
QY	2097	tttcacgcaggaattattttaaagcagcaggggaataatttaagatgaattaaagca	2156
Db	2734	TTTTCTTCAAGGTTTTCACATTGTGGCAAACTGTGAGAGATCTTCACCAAGCGCTCCGATA	2793
QY	2157	cttgcgccttcaaacgatgtlatatagaatactgccttggttctgtaagaagaagaaacag	2216
Db	2794	CATGTGTTTCGATTCATGCTGCAGACAGAGATGTGATGTGTTGATGACCTTAATAAAC	2853
QY	2217	cacagagccttgaacaacataltagccatagatatagcacaaaltcaaggagcaacag	2276
Db	2854	CACGTGAGCATTTACATACCGAGCGCTCCCAATATGCTGCTCAAGTTAAAGGGCTATAG	2913
QY	2277	ttccaggaattgggctcttaacacaaatgcaacttaacatgggacttcaagttacatctc	2336
Db	2914	TTTCCAGGCTTGCTGTATTTATTCACCGAATATATCAAAATCTTGGGACGAGAAATATTAATTC	2973
QY	2337	ctttatataatgatacgggagatgacatcccttgatctcgtatgtctgcagtaactgcgt	2396
Db	2974	CTTCATCTATGTGTGGCAACTATACACTGTACTCTTAGCAATTTGATACCATTCATTTGCAT	3033
QY	2397	gacaggaatgattgnaaacccgcagcaatgacatgacattgccaacaagaataagcaagaact	2456
Db	3034	AGCGAGGATTTGTGAATGATAATGTTATGTCTGACACAGCAGCTGAAGATAAGAAAGAACT	3093
QY	2457	taagcagatgcggaaagatagcaactgaacttggaaatatacgtactatagtgcat	2516
Db	3094	AGATAGTGCTGGAGAGTGTGCTACTATACCAATATGAAACCTTCCAAACCGTTGTTCTTT	3153
QY	2517	aacaaggaagaaagcccttcgagcaaaatgcatgaagatgacatgcagactcaacaaga	2576
Db	3154	GATCAGACGACAGAGATTGTAATATATGTATGCTCAGATTTGGCAGATGCCATATACGAAA	3213
QY	2577	taactcgaagaagaacacagattatggaagctgttatgcatcagccaatgacctatata	2636
Db	3214	CTCTTTTGAAGAAAGCACACATCTTTTGGATTTACATTTTCTTCAACCCAGGCAATGATGTA	3273
QY	2637	tttgcctatagcagcagggttttgatttggagcctatattcaactcaagctlgacgaatgac	2696
Db	3274	TTTTTCTATAGCTGAGATGTTTCCGGTTTTGGACCTCATTTGGTGACATTAACATCTCATAG	3333
QY	2697	cccaagagcagatgttcaatagattttcctgaacttcacataggaactatgagccatcgaa	2756
Db	3334	CTTTGAGAGATGTTCTGTATGATTTTTCACATGTGTGCTTTGGTGCATAGCGCGTGGGCA	3393
QY	2757	aagcgtcgttttgcctctgaatatcacaagccaatcggggcgtgcgcatctgtttgc	2816
Db	3394	AGTCAGATTCATTTGCTGCTGACTATCCCAAAACCCAAAATATATCAGACGCCCATCATCAT	3453
QY	2817	cttgttgaagaagaacccaataatagacagccgcagctcaagaaggaagaagccaagac	2876
Db	3454	GATATATGAAAAAACCCCTTTGATTAACGCTACACGACACAGCGAAGGCCATTAATCCGAAC	3513
QY	2877	atgtgaagaatattagatgttcgaaagctctctcttctcctacatcgtcgtccagaatgt	2936
Db	3514	ATTGGAAGGAATATGTCATATTGTGTAATTTGTATTTCAACTATCCACCCAGACCGGACAT	3573
QY	2937	tttcatctctcgtgcttaccctcagatltgagcgaagaaagacagtatgatttgtgg	2996

Accession	Gene	Protein	Length	Species	Reference
3574	CCCACTGCTTCAAGGACTGAGCTTGAGGTGAAGAGGGCCAGACGCTGGCTTGGTGG		3633		
QY 2997	gagcagcgctgtgagaaagcaacttcctgttcaacttcctcagaagaacttlaaccgct		3056		
Db 3634	CAGACGTGGCTGTGGGAAGAGCAGCAGTGGTCCAGTCTCTGAGGCGGTTCTAGACCCCTT		3693		
QY 3057	gcaaggaacagtctgtttttagtgbtgbatgcacaaagaatgaaftacagtggtccg		3116		
Db 3694	GGCAGGAGGAAGTGGCTGTGATGGCAAGAAATAAAGCGACGATGTTCAGTGGCTCGG		3753		
QY 3117	ttcccaatlagcaatcgttctctcaagaccctgtgtcttcaactgcacattgtctgaagaa		3176		
Db 3754	AGCAACACTGGGCATCGTGTCCAGAGACCACCTCTGTTGACTGCACATTGTCTGAAGAA		3813		
QY 3177	catgcacctatgttgcacacagccgtgtgtgtgcattatagatlagatcaagaagccgcaaa		3236		
Db 3814	CATTGCCATTGAGAGCAACACAGCCGGGTGGTGTCCAGGAAGAGATCGTAGGGCAGCAAA		3873		
QY 3237	tgcagcaaatatccattcttatttgaaggtctccctgagaaatacaacacaaagtgtg		3296		
Db 3874	GGAGGCCAACAATACATGGCTTCATCGAGTCACTGCTTAATTAATATACACTAAGATAGG		3933		
QY 3297	actgaagagagcaagcttctctgctgcgcagaaacaaagaactagctatttcaaggctct		3356		
Db 3934	AGACAAAGGAACCTCAGCTCTCTGTGGGCGCAGAAACAGCATTTGCCATAGCTCGGCCCT		3993		
QY 3357	tctccaaaaccaaaattttatttattgttttagtgaagccacttcaagccctcagtaagagac		3416		
Db 3994	TGTTAGACAGGCTCTAATATTTTGTCTTTTGGATGAAGCCACGTCAGCTTGAGATACGAAGA		4053		
QY 3417	tgaagaagtggtttcagcagctgccttgcataaagccagagccggaagagacatgcctagtgt		3476		
Db 4054	TGAAGAAGTTGTGCCAAGAGGCTTGAGCAAAAGCCAGAAAGGCGCCACTTCGATTGTGAT		4113		
QY 3477	cactcaacagctctcttgcattcagaatcagaacgcagattgatagtgtgtctgcacaatgaaa		3536		
Db 4114	TGCTCACCGCCTGTGCCACCATTCACAAAGAGCAGACTTAATAGTGGTTTCAGAAATGGCAG		4173		
QY 3537	gataaaggaagaagaactctcaagaagctcctgtagaagaatcagaacatatatttaagt		3596		
Db 4174	AGTCAAGGACGATGGCAGCATACACAGCACTGCTGGCAGAAAGGCACTAATTTTTCAT		4233		
QY 3597	agtgaatgcacagtcag 3613				
Db 4234	GGTCAGTGTCCAGGCTG 4250				
RESULT 9					
HUMMDR1					
LOCUS	HUMMDR1 4646 bp mRNA PRI 03-DEC-1999				
DEFINITION	Homo sapiens P-glycoprotein (PGY1) mRNA, complete cds.				
ACCESSION	M14758				
VERSION	M14758.1 GI:187468				
KEYWORDS	P-glycoprotein; drug resistance protein; transport protein.				
SOURCE	Human drug-resistant cell line KB-C2.5 cDNA to mRNA, clones lambda-NHRI10, 5, 1041.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B.				
JOURNAL	Internal duplication and homology with bacterial transport proteins				
MEDLINE	in the mdr1 (P-glycoprotein) gene from multidrug-resistant human				
REFERENCE	cells				
AUTHORS	Cell 47 (3), 381-389 (1986)				
TITLE	Ueda,K., Clark,D.P., Chen,C.J., Roninson,I.B., Gottesman,M.M. and				
	Pastan,I.				
	The human multidrug resistance (mdr1) gene. cDNA cloning and				
	transcription initiation				







OY	3237	tgcagcaaatatccattcttttattgaaggttcctcctgagaatatcacacaacaaagttag	3296
Db	3874	GGAGGCCAACATCAATGCGCTTCATCGAGTCACTGGCTTAATAATATGACCTAAAGTAGG	3933
OY	3297	actgaagaagagacagccttcctggcgccgaagaacaaagacagcatatgcaaggccttc	3356
Db	3934	AGACAAGAAGACTCAGCTCTTGCTGGCCGGAACACACCATTGGCANTAGCTGTGCCTT	3993
OY	3357	lctccaaaaccacaataattltatgtgatlgaggccacttlcagccctcgcataatgacag	3416
Db	3994	TGTTAGACAGCCTCATATTTTGTCTTTTGGATGAGGCACAGCTCAGCTGTGATACGAAGA	4053
OY	3417	tgaagaagtggtgtccagcatgaccttgataaagccagagcgaggaaagaaatgacctgtgt	3476
Db	4054	TGAAAAAGTTGTTCACAGAAGCCCTGGACAAAACCCAGAGAAGGCCGACCTGCATTGTGAT	4113
OY	3477	cactcaagagcctcttgcaattcagaagccaatttgttaagtggtcttgcacaaatggaaa	3536
Db	4114	TGCTCACCCCTGTCCACCATCCAGATGCACACTTAATATGTGTGTTACAAATGGCAG	4173
OY	3537	gataaagagacaagaatcatcatalcaagaagctctctgagaaaatcgagacatatattaagt	3596
Db	4174	ACTCAAGAGAGCATGGCCACCATCAGCAGCTGCTGGCACAGAAAGGCATATTITTTCAAT	4233
OY	3597	agtgaatgcacagtcag 3613	
Db	4234	GSTCAGTGTCCAGGCTG 4250	
RESULT	10		
LOCUS	108557	4669 bp	PAT 02-DEC-1994
DEFINITION	Sequence 3 from Patent WO 8705943.		
ACCESSION	108557		
VERSION	I08557.1 GI:588735		
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 4669)		
TITLE	Romanson,I.B., Pastan,I.H. and Gottesman,M.M.		
JOURNAL	COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS		
FEATURES	Patent: WO 8705943-A 3 08-OCT-1987;		
source	Location/Qualifiers 1..4669 /organism="unknown"		
BASE COUNT	1394 a 892 c 1129 g 1254 t		
ORIGIN			
Query Match	36.3% Score 1316.2; DB 6; Length 4669;		
Best Local Similarity	62.3%; Pred. No. 3.1e-266;		
Matches 2178; Conservative	0; Mismatches 1268; Indels 51; Gaps		
OY	162	caggttaccctgatatctatgttggaataggttgcctgcctgattttgtgttacataca 221	
Db	760	CAGGATATGCTATTATTAAGTGGAAATGGTGTGGGCGTGGTGTGCTTACTTTCA 819	
OY	222	gatttcctcttggtatattactatgcagcacgacagacacaaagagattcgaaaaagtttt 281	
Db	820	GGTTTCATTTTGGTGTGCTGCAGCTGGAGACAAATATACAAAAATTAGAAAACAGTTTTT 879	
OY	282	tatttcagtttgggacagagacatcggtcgtgtttgatacgttgacatcgttgygaacttaa 341	
Db	880	TGATGTCTAATAATGCGACAGGAATAGGCTGTGTGATGTGACAGATGTTGGGAGCTTAA 939	
OY	342	cactcgcataagacatcagac---aaatcagtgatgttatggagataaagtctct 398	
Db	940	CATCCGACTTACAGATGATGCTCTTAAGATTATATGAAGTATTGTGGACAAAAATTGGAAT 999	
OY	399	gttgtttcaaacatgtctacttttgcatttgacctgacctgagttgttggtaaggcttg 458	

Db	1000	GTCTCTTACGTCATGGCAACATTTTTCACGTGGGTTTATAGTAGATTTACACGTGGTTG	1053
QY	459	gaactaacccatgtagctcatcacagctccctctcttaaatgagctcagcgcgacatg	518
Db	1060	GAGCGTAAACCTTGATTTTGGGCATCAATCCTGTTTGGACGTGACGTGCTG	1119
QY	519	tcttagatgtagctcatcattgaccagtaagaattaaigtccctattccaagctgggc	578
Db	1120	GGCAAGATTAATCTTCACTTACTTACTGATTAAGAACCTTAAAGGTATGCAAAAGCGTAGC	1179
QY	579	tgtagcgaagaagctctgtcatcatccgaagcatagcctcttagggcccaagaaa	638
Db	1180	AGTAGCTGAAGAGGCTTGGACGCAATTAACACTTGATTTGATTTGGACGCAAAAGAA	1239
QY	639	agaactccaagagatcacagatcccaagaatgcaagaatttlygataaaaaagagc	698
Db	1240	AGAACTTCAAAAGTACACAAAATAATTAGAGAGAGCTAAAGAAATTGGGATTAAGCAAGC	1299
QY	699	tatagctcaaaagtgctccttggtgctgtgtgactctcttatgatgaaactatgagct	758
Db	1300	TATTACAGCCAAATTTCTTATAGTGGCTGTTCTCGTGATCTATAGCATCTTATGCTCT	1359
QY	759	tgcttttltgtatggaaccccttgattctcttaatlgagaacctgatatacactcgagc	818
Db	1360	GGCCTTCTGGTATGGGACCACTTGGCTCTCCAGGGG-----ATATTTATTGGACA	1413
QY	819	tgcttgcgttttcttcttgatgtaatccatagcagtaattgcatlgagcgagctcc	878
Db	1414	AGTACTCACTGTANTCTTTCTGTATTAAATGGGGCTTTTAGTGTGGACAGCATCTCC	1473
QY	879	tcaacttgaacctgcgcaatagcccgagagctgtcccttcataatttccagttatga	938
Db	1474	AAGCATTTGAACATTTTGCATTAATGAAAGAGAGAGAGCTTTATGAAATCTTCAATATTTGA	1533
QY	939	taagaaacccagatataacttcttccacagctgataataacctgaatccatagaag	998
Db	1534	TAAATAGCAAGTATTTCACAGCTATTGGAAGAGTGGGCAAAACGATATATTATTAAGG	1593
QY	999	aacgtggaatttbaaaatgcttcttcaattatccatacaagaccatcatcaagattc	1058
Db	1594	AAATTTGGAATTCGAANAATGTCTCACTTCATTTACCATCTCGAAAAAGATTAAATCTT	1653
QY	1059	gaaagtlctgaatctcagaaatlaagtlctgagagacagctgcctgtgtcgtlctcaatg	1118
Db	1654	GAAAGGCTGAACTGGAAGGTGCAGAGTGGGACAGAGGTGGCCCTGTTGAAACAGTGG	1713
QY	1119	cagtgaggaagatcagtagtcaagctcttcgcaaggttatatgatccgatagagctt	1178
Db	1714	CTGTGGGAAGGCACAACAGTCCAGCTGATGAGAGAGGCTCTATAGACCCACAGAGGGAT	1773
QY	1179	tatactgtgatataagaatcacatcaaggtttaaattgcygcacatcatcgcagacatat	1238
Db	1774	GGTCAGTGTATATGACAGGATATTAGGACCATTAATGTAAAGTTTCTACGGGAATCAT	1833
QY	1239	tggagtggtltagtcaagaagcctgttltgttcggagaccacatcaytaacatatcaagta	1298
Db	1834	TGCTGTGATGATCAGGACACCTGATTTGTTGCCACCAAGATAGCTGAAMAATCTCGCTA	1893
QY	1299	tggacagagatgattgtgactatgaagagatlgagagagacgacgaagggaaagcaatgcta	1358
Db	1894	TGGCGCTGAATAATCTCACATGTGTGATGTGAAGAGCTGTCAAGGAAGCCCAATAGCTTA	1953
QY	1359	tgattttatcattgagttctcttaataattatataattgtgtaggggaaaaagagctca	1418
Db	1954	TGACTTTATCTGAAGACTGCTCTAATAATTGACACCTGTGTGGAGAGAGAGGGGCCA	2013
QY	1419	aatgattgagggagaaacagagatcgaattgtcgtgtgacttagttcgaaaccccaa	1478
Db	2014	GTTTGAAGTGTGGGACAGAAAGAGAGATCGCCATTGGACGAGCCCTGTTCCGCAACCCAA	2073
QY	1479	gattctgattttagatgaagctacgtctgcctctgattcaagaagcaagtacgtcttca	1538
Db	2074	GATCTCTCTGCTGTGATAGGCAACGTCAAGCTTGGACACAGAAAGGAAGCAGTGGTTCA	2133

QY	1539	agccgcacctggagaaagcgagcaaaagtcgcgacataacatcgtgtgtagcaccagcattc	15988
Db	2134	GGTGGCTCTGTGATTAAGGCCAGAAAAGGTGGACACACATTGTGATAGCTCATCGTTTGTG	219393
QY	1599	tactatcgaagltgcagatttgaattgtgaacctaaagaaatggaaatgcttgcgttgcgaaag	16588
Db	2194	TACAGTTCGTAAAGCTACAGCTCATCGCTGATTCGATGATGAGAGTCATGTTGGAGAAAGG	22533
QY	1659	agccacatgcctgcgaactaaatgcaaaacgaggtctatattatctacttgcgtgcgaagaa	17188
Db	2254	AAATCATGATGAACCTCATGAAAGAGAAAGCAATTTACTTCAAACTGTGTCACATGCGACAC	23133
QY	1719	tattaaaaagcctgcgaagaacagaatggacaaatgacatctta-----	17628
Db	2314	AGCAGGAAATGAAGTTGAAATTAATAAAATGCACTGATGATCCAAAAGTGAAATTGATGC	23733
QY	1763	-ctgaagaagaagaacaaactactcctctgcacactcgtgaagagcatca-----	18108
Db	2374	CTTGGAAATGCTTCAATATGATTCAGAGTCCAGTCTAATTAAGAAAAGTCAATCGTAG	24333
QY	1811	-----agtcgagactctatgcagaagctgcgaatccaccacaatctaagagat	18598
Db	2434	GAGTGTCCGTGATACAAAGCCCAACAGAAAGCTTACTACCAAGAGGCTCTGGATGA	249393
QY	1860	aagcttcctcgaagtcctctatlaaaaaatttaaaqttaaacaagcctgcgaatgacctt	19198
Db	2494	AAGTATACCTCCAGTTTCTTTTGGAGGATTAATGAAGCTAAATTTAACTGAATGGCTTTA	255393
QY	1920	tgtgtctcgtggagacattgctctcgttctcaaatggaaatgctgtccagatatttccat	19798
Db	2554	TTTTGTGTTGGATATTTTGTGCCATTTAATATGAGGCGTCGAAACGACGATTTTGCAAT	26133
QY	1980	catctttgcaaaaatatacaacatgltt---ggaaataatgaataaccacaatlaaaga	20368
Db	2614	AATATTTTTCAAAGATTAATAGGGTTTTTACAGAAATGATGATCTCGAAACGAAACGACA	26733
QY	2037	tgaatgcgaatatatacccatglatctgcatattgggtatattgttcgttgcagtta	20968
Db	2674	GAAATGTAACCTGTTTTCTACCTATTGTTTCTAGCCCTTGGAAATTAATTTCTTTAATTAAT	273393
QY	2097	tttcaatgcaggaattatatttacgcagagacgagggagaaatttlaaagatgcgaatlaaga	21568
Db	2734	TTTTCTTCAAGGTTTTTCACATTTTGGCAACGCTGGAGAGATCCACACAAAGCGGCTCCGATA	279393
QY	2157	cttgcgctctaaagccaatgltatataatgaatatatgcctgcgttgcgtatgaatgaagaac	22168
Db	2794	CATGCTTTTCCGATCCATCGCTCGACAGAGATGATGTTGTTGATGACCCCTTAACAAACAC	285393
QY	2217	caagagagagcttcaacaacaaatctagccatagatatagacaaatcttaagagcgaaacg	22768
Db	2854	CACCTGGACATTGACTACACAGGCTCGCCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGG	29133
QY	2277	ttccaagatctgcgtctcttaacacaaaatgcacaaatcaacatgcggagacttcagttatcatlc	23368
Db	2914	TTTCCAGGCTTGCATGATTAATACCCAGAAATTAAGAAATCTTGGAGACAGAAATTAATATAC	29733
QY	2337	cttataatatgabatvggagatgcacatctcgtatctgagatattgcctcagtaactgcgcgt	23968
Db	2974	CTTCACTATGCTTGGGCAACATAACCTGTTACTCTTACCAATTAATGACCATCATATCAAT	303393
QY	2397	gacaggaatgcgaatgaacccgcagcaatgcgtgcgatttgcacaacaagaatgaagaact	24568
Db	3034	AGCAGGAGTTGTTGAAATTGAAAAATGTTGCTGGACAAGCACTGAAAGATTAAGAAAGAACT	309393
QY	2457	taagcatgctgcggaagaatgaacaatgcgaatgcgttgcgaatatgcactagatgcatt	25168
Db	3094	AGAAAGTGTGGGGAAGATGCCACTGAACCAATTAAGAAACTCCGACCGTGTGTTCTTT	315393
QY	2517	aacaaaggaaaagccttcgagcaaatglatgaagaatgcttcagactcaacacagaaa	25768
Db	3154	GATCTACAGACAGAAATTTGAACATATGATGCTCGAGATTTGACAGTTCACATGACCTATAGAA	32133

OY	2577	tactctgaagaagcacacagattatitggaagcgttctatcgaatcgatgccttatata	2636
Db	3214	CTCTTTGAGGAAACACACACTCTTTGGAAATTACATTTTCTTCCACCCAGGCAAATATGTA	3273
OY	2637	tttgcacatgcagcagggcttcgatitctggagccttatctaacaagcctgcgagatgc	2696
Db	3274	TTTTTCTATCTGAGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACAATAACTATATG	3333
OY	2697	cccaaggagcattgcatatcttaactgcgaattgacatatggaactatgcccacgaaa	2756
Db	3334	CTTTGAGAGATTTGTTGATATTTTACGATGTTGCTTTGGTGGCCATGGCCGTGGGCA	3393
OY	2757	aagctcgttttgcctctctaattatccaagaacaaatcgggggctgtgcgaatctgttc	2816
Db	3394	AGTCAGTTCAATTTCTCTCTCAATGCGCAAGCCAAATAATATCAGACGCCACATCATAT	3453
OY	2817	ctgtctggaanaagaaaccaataatagacagccgcagtcgaagaagggaanaagccgaac	2876
Db	3454	GATCATTTGAAAAAACCCCTTTGATTTGACACTACACAGGAAAGGCGCTAATGCGACAC	3513
OY	2877	atgtgaagggaatttagagttctcgagaagctctcttctctatccaatgtgcacagatgt	2936
Db	3514	ATTGGAAGGAATCTCATTTTGGTGAAGTTGATTTCACTATCCACCCGACCGAGAT	3573
OY	2937	tttaacctccgtgcctatccctcaagatctgagcaggaagaagaagatgcattgtggtg	2996
Db	3574	CCCACTGCTTCAGGAGACTGAGCCTGGAGGTGAAGAAGGCGCACCGCTGTGTGTGG	3633
OY	2997	gagcagcgcgtgtgggaanaagcaactctgttcaactctgcagagacttatatgacccgt	3056
Db	3634	CAGACAGTGGCTGTGGGAAGACACAGTGGTCCAGCTCTCGAGCGGTTCTACGACCCTT	3693
OY	3057	gcaaggacaagtgtctgttgaatgtgtgtagtgaanaagaattgaattgaacagtgctcg	3116
Db	3694	GGCAGGGAATGCTGCTTATGACAAAGAAATAAGACATGAATGTTCAGTGGCTCG	3733
OY	3117	ttcccaaatagcaatctgtccctaagaagccctgtgccttcaactcagcaattgcgaa	3176
Db	3754	AGCACACTGGGCATCGTGTGCCAGGAGCCCATCTGTTTGACTGGACGATGTGTGAGA	3813
OY	3177	catgcgtctatgttgaacaagccgltgtgtgcattagatgagatccaagaagccgcgaa	3236
Db	3814	CATTGCTTATGAGAACAAACGCGGGGTGTGTACAGGAAGAGATCTGTAGGGCAGCAAA	3873
OY	3237	tgcaagcaaatctcatctcttctaagaagtctccctcggaataacaacaacaagtgtg	3296
Db	3874	GGAGGCCAACATTAATGCTTCACTCAGATCACGTGCTTAATAATATAGACATTAAGTAG	3933
OY	3297	actgaagagcacagctttctggcggcgcaaaaacaagaactagtatgtgaagggctct	3356
Db	3934	AGACAAAGGAATCAGCTCTCTGTGGTGGCCGGAACAACCAATTTGCCATAGTCGTGCGCT	3993
OY	3357	ttctcaaaaaccceaaatttlatitgtgtgtatgtgtagtgcacatcttaagccctcgataatgcag	3416
Db	3994	TGTTTAGACAGCTTATATTTTGTCTTTGGATGAAGCCAGCTCAGCTTGTGATACGAAGA	4053
OY	3417	tgaaagagttgttcaagaatcccttgataaagccaggaacgggaaggaacatgtcctagtgt	3476
Db	4054	TGAAGAAGTTGTCAGAAAGCCCTGGACAAAGCCAGAGAAGGCGGCACTTGATTTGAT	4113
OY	3477	cactcaacaggtctctctgcaattcaagaacgagatttgaatgtgtctgtgcacaatgga	3536
Db	4114	TGCTTCACCGCTGTGCCACATTCAGAAATGCACTTAATATAGTGTGTTTCACAAAAGGAG	4173
OY	3537	gataaagaacaagaactatcaagaagcctcgagaataatcgagacataatctttaaagt	3596
Db	4174	AGTCAGAGAGATGACACGATCATCAGACGTGCTGGACAGAAAGGATCTATTTTCAAT	4233
OY	3597	agtgaaatgcacagtctag 3613	
Db	4234	GGTCAGTGTCCAGGCTG 4250	



[illegible]

Db	5026	CAGCAGTGGCTGTGGGGAAGAGCAGCACTGGGTCCAGCTCCTTGAGAGGGGTTCTACAGACCCCTT	5085
Qy	3057	gcaagagaaagatgctgttttgaatggtgtgatagtcaaaagaattgaaatgacagtggtctcg	3116
Db	5086	GGCAGGAAAGTGGCTGCTTGTATGGCAAAAGAAATAAAGCAGCACTGAATGTTTCAGTGGCTCCG	5145
Qy	3117	ttcccaaatgcaatggtttctctaagaagctgtgacctcaacatcacaattgtctggaa	3176
Db	5146	AGCACACCTGGGCAATCTGTGTCACAGAGGCCATCTCGTTTGACTGCACCATTTGCTGAGAA	5205
Qy	3177	catgcctatgtgtacacaacagccgtctggtgtccattagaatgaaatcaagaagacgcgcgaa	3236
Db	5206	CATTGCTCATGTGAGACAAACAGCCGGGTGGTGTACAGGAAGAGATCTGTGAGGCAAGCAAA	5265
Qy	3237	tgccagcaaatatccatctctttatagaagtlctcccctgaaatacaacacaagaattgg	3296
Db	5266	GGAGGCCAACATATCATCTGCTTCATCGATGACAGTCAGCTCTAATAATATATGACTAAATGAG	5325
Qy	3297	actgaagagagcaagctttcttggtggtggtccagaacaagaacacagctattgtcaagggctc	3356
Db	5326	AGACAAAGAACTAGAGCTGCTCTGGTGCCAGAAACCAATATGCTATAGCTCGTGCCCT	5385
Qy	3357	tctccaaaaccceaatttattgtgttgaatgagcgcactcaagccctcagataatgagac	3416
Db	5386	TGTTTAGACAGCCTATATTTTGGCTTTTGATGATGAAGCCACAGCTGCTGTGATACGAAG	5445
Qy	3417	tgagaagttgttltacgaatgcctctgaataagaacgaagacggaagaaagacatgctagtgt	3476
Db	5446	TGAAGAAGTTGTCCAGAGAGCCGTGGACAAAGCCAGAGAAGGCCGACCTGCAATTTGTAT	5505
Qy	3477	cactcaacagctctctcgtcaattcaatgaacgcagattgtgagtgtctgcacaaatgaaa	3536
Db	5506	TGCTTCACCGCTGTGCCACATCCAGATGAGACTTAAATAGTGATGCTTTCAGAAATGGCAG	5565
Qy	3537	gataaaggaaacaagaacatcatcaagaagctccgtgaaatcgagacatatatttaagt	3596
Db	5566	AGTCAGAGAGCATGCGACGCACTGACGACGCTGTGCACAGAAAGGACATCTATTTTCAAT	5625
Qy	3597	agtgaatgcagacgtcag	3613
Db	5626	GGTCAGTGTCCAGGCTG	5642
RESULT	12		
LOCUS	ARO28672		
DEFINITION	Sequence 6 from patent US 5858744.	DNA	29-SEP-1999
ACCESSION	ARO28672		
VERSION	ARO28672.1		
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 9318)		
AUTHORS	Baum,C., Stocking-Harbers,C. and Osterlag,W.		
TITLE	Retroviral vector hybrids and the use thereof for gene transfer		
JOURNAL	Patent: US 5858744-A 6 12-JUN-1999;		
FEATURES	Location/Qualifiers		
source	1..9318		
BASE COUNT	2377 a 2217 c 2364 g 2358 t	2 others	
ORIGIN			
Query Match	36.3%;	Score 1316.2;	DB 6; Length 9318;
Best Local Similarity	62.3%;	Pred. No.3.1e-266;	
Matches 2178;	Conservative	0; Mismatches 1268;	Indels 51; Gaps
Qy	162	caaggtgacctgtatattgttgaataggtgtgtgcttgcttgatttggttacataca	221
Db	2111	CAGGATATGCCATTATTTATTCACGTGGAATATGCTCGCGGGTGTGCTTGCTGCTTACATTTCA	2170
Qy	222	gatttcctgttgattataactgcagcagcacagaccaaagagattcgaaaacagtltt	281



Db	2171	GGTTTCATTTTGGTGCCCTGGCAGCTGAGACAAATACCAAAATTTAGAAAACACTTTT	2230
Oy	282	tcattcagtttttgcacagagacatcggcttggtttgaatactgtgacatcgttgaacttaa	341
Db	2231	TCATGCTTAATATGGCAGACAGGAGATAGGCTGTTTGAATGTGCACAGATGTTGGGACCTTAA	2290
Oy	342	cacttcgcatacagacatgtac---aaatcagtgatglatgtgagataaagtctctc	398
Db	2291	CACCGGACTTTCACATATATGCTCTCTAAGATTAATGAAGTAAATTTGGTGCACAAATATGGAT	2350
Oy	399	gttgcttcaaaaacatgtcctactcttttcgatattgaccgagatgtgttgggtgaaggctg	458
Db	2351	GTTCTTTGACGCAATGGCAACATTTTTCATCTGGGTTTATAGTAGGATTTACACGNGTGG	2410
Oy	459	gaactccaccttagtgactcatalccacgctccctcccttaataatggtctcagcgacagatg	518
Db	2411	GAAGCTAACCCCTTGATGTGATTTTGGGCATCAAGCTCCTGTCTTTGGACGTGTACGCTGTCTG	2470
Oy	519	ttctcagatgtgtcatctcattgacacagtaagaaataagtgccattlccaaagctgggac	578
Db	2471	GGCAAAAGTACTATCTTCATTTACTGTATAAAGAACCTTTAGGCTATGCAAAAGCTGGAGC	2530
Oy	579	tgtgccaagaagaatctctgtccatcaatccgaacagtcataagcttaaggccccagaa	638
Db	2531	AGTACCTGAAAGGCTCTTTGGCAGCAATTTAGAACTTGATTTGCAATTTGGAGCAAAAGAA	2590
Oy	639	agaactc aaagatatacacagaatccaaagaatgcaagaagatlttgcataaaaaagac	698
Db	2591	AGAAGCTGAAGGTACAAACAAATAATTAGAAAGAGCTAAAGAAATTTGGGATTAAGAAAGC	2650
Oy	699	tatagcttcaaaagtgtctctgtgtgtcgtgtgactcttataagatggaactatgact	758
Db	2651	TATTACACCAATTTTCTATAGGTGCTGCTTCCCTGCGATCTATGCACTTATATGCTCT	2710
Oy	759	tgcttttggatagaaaccccttgatcttcttaatggaagacctgatatcacatcgagac	818
Db	2711	GGCCTTCTGGTATGGGACACACTTTGGTCTCTGAGGGGA-----ATATCTTATTGGACA	2764
Oy	819	tgcttttgcttcttcttctttagtgaatccatlaagcaatlttgcattggaacagacatcc	878
Db	2765	AGTACTCAGCTGTATTTCTTTCTGTATTAATTTGGGGCTTTTATGTTTGGAACAGCAATCTCC	2824
Oy	879	tcaacttggaaacctgcgaatagccgagagagctgccttccatatttccagtatatga	938
Db	2835	AANGATGGAACATTTTCCAAATGCAGAGAGGACGACTTATGAAATCTTCAGATATATTTGA	2884
Oy	939	taagaaacccagatagatatcatttccacagcttgatataaactcgtatcatcatgaagg	998
Db	2885	TAAATAGCCAAAGTATTCACAGCTATTGTGAAGAGTGGGCAAAACAGATATATTTAAGG	2944
Oy	999	aactgtggaatttlaaaaatgttcttcttcaaatlccatcaagaacatcatcaagaatct	1058
Db	2945	AAATTTGGAATTCAGAAATGTTCACCTTCAGATTACCATCTCCAAAAAAGAAATTAAGACTTT	3004
Oy	1059	gaaaggtcctgaatcctcagaatlaagtcctgagagagacagctgcgctgtgtcgtcctaag	1118
Db	3005	GAAAGGCTCGAACTGTGAAGGTGCAGAGTGGGCGAGAGGTGGCCCTGTGGTGAACACGTGG	3064
Oy	1119	cagtgaggaaagatccggttaatccagcttctgcagagagtatatgtatcggatagatgctt	1178
Db	3065	CTGTGGGAGACGCAACACAGTCCAGCTGATGACGAGAGGCTCTATTGACCCACAGAGGGGAT	3124
Oy	1179	tatcatgttgatbagaatagatacatcagagctttaaattgycgcaatltcagagacat	1238
Db	3125	GGTCAGGCTTGATGACAGAGATATTATAGACCAATTAATGTAAAGGTTTCTTACGGGAATATAT	3184
Oy	1239	tggagtggttagtcaagaagcgttcttgtctcggagaccacatcagtaaacatataagta	1298
Db	3185	TGCTGTGTGTGATGAGGAACCTGTATTTGTTGGCACACAGATAGCTGAAACATTTGCGTA	3244
Oy	1299	tggagcagatgatatgtactgtatgaagatatggaagacagcaagggaagcaaatgtcgta	1358

Db	3245	TTGGCCGTTAAAAATGTCACACATGATGATGAGATTGGAAAGAAAGCTGTCCAGGAAGCAATAGCTTA	3304
Qy	1359	tgattcttcaatcttgagatttccctaataatttaatacatctgtagagggaaaaagagctca	1418
Db	3305	TTACTTTATTCATGAAGACGCTCCTATAATTGTGACACCGCTGGTTGGAGAGAGGGGCCCA	3364
Qy	1419	aatgagtgagggcagaacaaacagagatctgcaattctctgtgtccctagtctgaaccccaa	1478
Db	3365	GTTCAGTGTGGTGGGCGAAGACGAGAGAGATCGCCATTGGACGTCGCCCTGTTCGCAACCCAA	3424
Qy	1479	gattctgatttagatagagctacgtctgcccggatctgaagaacagtgccgctgttca	1538
Db	3425	GATCTCTCTGCTGGATAGGGCCAGTCAGCTTTGGACACAGAAAGGAGACGCTGTGTCA	3484
Qy	1539	agctgcacttgagagaagcgagcaaaagctcgagactacaatctgtgttagcacaccgacttc	1598
Db	3485	GGTGGCTCTGGATTAAGGCGAAGAAAGGTGGCACACCATTTGTGATAGCTCATCTGTTTGTG	3544
Qy	1599	tactatctgaaagtgcagatttgattgtgacctaaagatgtaatgctgtgcgagaaag	1658
Db	3545	TACACTTTGTGATATCTGACGTCATCGCTGGTTTCGATGTAGTGAGCATCTGTGTGAGAAAGG	3604
Qy	1659	agcaaatctgaactaaatgcaaaacgaggtctatatcttcaacttggatgtagtcacaga	1718
Db	3605	AAATCATATATACATCATGAAGAGAAAGGCATTTTCTTCAAACTGTGTCAATCTCAAC	3664
Qy	1719	tattaaaaagctgaatgaacaagatgtagtaatgatactca-----	1762
Db	3665	AGCAGAAATGATGTTGATTTAGAAAATGACGATGATGAAATCCAAAGAAATTGATGC	3724
Qy	1765	-ctgaagaagaaccacactcactctctgcaactctgtgaagagca-----	1810
Db	3725	CTTGGAAATGTCTTCAATGATGTTTCAGAGTCACCTCTAAATAAGAAAGATCACTGTAG	3784
Qy	1811	-----agtcagaattctatgcaagctgtgagatccaccacatccaaagat	1859
Db	3785	GAGTCTCCGTGGATACAAAGCCCAAGCACAAAGCTTACTACCAAGAAGCTCTGGATGA	3844
Qy	1860	aagctctcctaagctcctctatbaaaactttaaagttaaaacaacgctgaaigtccctt	1919
Db	3845	AAGTATACCTCCACTTTCCTTTGGAGAGATTATGAAGCTAAATTTTAACGTAAATGGCCCTTA	3904
Qy	1920	tgtgtctctgggagacatggtctctgttctaaatggaaactgtcatccagtaatttcacat	1979
Db	3905	TTTTTGTGTTGGTATATTTTGTGGCATTTAAATGAGAGCCCTGCACACAGCATTTTGCAT	3964
Qy	1980	catcttgcaaaaatttaaacatgltt---ggaaaatagtaaaacaaatgaaga	2036
Db	3965	AATATTTTCAAAAGATTATAGGGTTTTTACAAAGAAATGATGATCCTGAACAAACGACA	4024
Qy	2037	tgatgcagaaaattatcatgatatctgcattgtgggtgatattgcttgttcaagta	2096
Db	4025	GAAATGTAACCTGTGTTTTCACATTTGTTTCTAGCCCTTGGAAATTATTTCTTTATTACATT	4084
Qy	2097	tttcaagaggaattatttaacgagcagagcagggaatttaacgatgagattaagaca	2156
Db	4085	TTTTCTTCAGGGTTTTCACATTTGGCAAAAGCTGAGAGATCTCCACCAAGCGCTCCGATA	4144
Qy	2157	cttgccctcaaacgcatgtatatacgaatctgctcgtgttggatgaagaagaaacag	2216
Db	4145	CATGGTTTTCCGATCCCATAGCTCAACACGAGATGTGAGTTGGTTGATGACCCTTAAACAC	4204
Qy	2217	cacagagagcttgcaacaacatattgacatagatagacacaacttaaggagcaacag	2276
Db	4205	CACGTGACCATTTAGCTACAGGCTCCGCATATGATGCTGTCAAGTTAAAGGGCGTATAGG	4264
Qy	2277	ttcagaatttgcgltcttaacacaaaatgcaactaacatcggacttcaagtatcatctc	2336
Db	4265	TTTCAGGCTTCTGTATTATCCCAAGATATAGCAAAATCTTGGAGACGATTAATATATATC	4324
Qy	2337	cttatataatgtagtggaatgacatctctgattctgagatctgcagtaactgtgcgt	2396
Db	4325	CTTATCATGATGTTGGCCAACTTAACGCTTACTCTTTCAGCAATTTATTAACCATCATTTAGAT	4384



OY	2397	gagggagatgatactgaaacccgacgaacatgacatgacttctgcacaacaaagtataacgaact	24585
Db	4385	AGCGAGGAGCTGTGTGAATAATGTTTGTCTGGACAGACACTGAAAGATAAAGAAAGAACT	44444
OY	2457	taagcatgctcggaaagaatacaactgaagctcttggagaataacgtactactatgctcaat	25168
Db	4445	AGAAAGGTCTGGGAAGATGCCCTACTGAGCAATAGAAAACCTCCGACCGTGTTCCTTT	4504
OY	2517	aacaaggaaanaagccttcgagcaaatgtatgaagaagatgcttcagactcaacacaga	2576
Db	4505	GACTCAGAGACAGAAATTTGAACATATATATGCTCAGATGATTGGAGTATACATACAGAAA	4564
OY	2577	tactctgaagaagaacacagatatcttggaaagctgtatgcatctcagccatgcttata	2636
Db	4565	CTCTTTAGAGAAAGCACACATCTTTGGAAATATACATTTTCCTTACCCAGCCAAATGATGA	4624
OY	2637	tttgcctatgacgacgaggttctgattttggagccttaatactcaagctcgaacgaatgac	2656
Db	4625	TTTTTCCATATGCTGGATGTTCGGTTTGGAGCTTACTTTGGGCACATAAACTCATATAG	4684
OY	2697	cccaaggaggaatctcaatgatttctaacctcaatctgacatactggagcatalgacatcgaa	2766
Db	4685	CTTTGAGAGATGTTCTGTATATTTTCACGTGTGTCTTTGGTGCATATGGCGTGTGGGCA	4744
OY	2757	aagcgtcgttctggtcctcgaatatctcaagaagcacaatcggggctgcgcacgtctgttc	2816
Db	4745	AGTCAGTTCAATTTGCTGCTCCTACTATCTCCAAAGCAAAATATACGACGCCCATCATCAT	4804
OY	2817	ctgtctgaaagaagaaccaaataataagccggcggatcaagaaggaaanaagccagaac	2876
Db	4805	GATCATTTGAAAAAACCCTTTTGATTAACAGCTACACACACGAGAGGCTTATATGCCAACAC	4864
OY	2877	atgctgaaggaaatctagagattctgagaagctctctctctctctcaatctgcgcagaagct	2936
Db	4865	ATTGGAAAGAAATGTCACTATTGGTGAAGTATTTCAACATATCCACCCAGCCGGAAT	4924
OY	2937	tttaactccctcggtgcttaccctcagatctgaagcagaagaagacagatgacattgttgg	2996
Db	4925	CCCACTGCTTACAGGACTGTGACCTGGAGGGAAGAGGGCCAGACGTGGCTCTGTGGGG	4984
OY	2997	gagagagagctgtggggaagaagcctctgttcaactcttcgacgaacttatgacccgt	3056
Db	4985	CAGAGATGGCTGTGGGAAGACACAGTGGTCAAGCTCTCTGGAGCGGTCTTACGACCCTT	5044
OY	3057	gcaagagcaagctgctgtttaaattgttgatgtaagaagaattgaatgtacagctgtctcg	3116
Db	5045	GGCAGGGAAGCTGCTGTTATATGGCAAGAATAAAGCACATGTAATTTTCACTGGCTCCG	5104
OY	3117	ttcccaaatagcaatcgcttctcotaagaagcctgtctcttcaacttcgaacatctgctgaag	3176
Db	5105	AGCACACCTGGGCAATCGTGTGCCAGAGCCACATCGTTTGACTGCAGATTGTGTGAATA	5164
OY	3177	catcgctctatgtgtgaacaaagccggtgtgtgtccattagatgtagatcaagaagccgca	3236
Db	5165	CATTGCTTATGGAGACAAACGCCGGGTGTGTACAGGAAGATGCTGAGGGCAGCATA	5224
OY	3237	tgacgaatatcatcatctcttcttgaaggtctcccttggaaatacaacaacaaactgttg	3296
Db	5225	GGAGGCCCAACATACATGCTTCTCATGATGTCATGCTCTTAATAATATPAGCACTAAAGTAGG	5284
OY	3297	actgaagaagacagcttcttgcgcgcgaagaacaaagactcagatatgtcaaggtcct	3356
Db	5285	AGACAAAGAACTACAGCTCTCTGTGGCCAGAAACAAAGCAATTTGCCATATAGTCGTGCCCT	5344
OY	3357	tctccaanaaccacaatttatactgttgaatgaaggccaactcagccctcgataatgacag	3416
Db	5345	TGTTTAGACACCTCATATTTTGTGTTGGATGTAAGCCACGTACGTCTGATATACAGAAAG	5404
OY	3417	tgaaaggtgtgttcaagaatccctctgaataagaacgaagcaggaagaagaatcctagttgt	3476
Db	5405	TGAANAAGGTGTCCAGAGAACCTTGACACAAAGCCAGAGAGGCCGCACCTGTGATTTGAT	5464

QY	3477	caatcaagagctcttcgcatctcaaacgcaatttgaatgattcgtgcacaaatggaaa	3536
Db	5465	tgcctacacccctgtccaccatccagatcagacttaataatgctgttcacaaatggcag	5524
QY	3537	gataaagagacaagaacatccatcaagaagctctctgagaatctcagacataatcttaagtt	3596
Db	5525	aattcaagagacatggcagaccatccagagctgctggcgcacagaagaagcattatTTTTCAAT	5584
QY	3597	agtgaatgcacagtcag	3613
Db	5585	gctcagctgtccacagcctg	5601
RESULT	13		
E02326			
LOCUS	E02326	4378 bp	RNA
DEFINITION	Multi drug resistance relating gene derived from human normal cells		PAT 29-SEP-1997
ACCESSION	E02326		
VERSION	E02326.1	GI:2170561	
KEYWORDS	JP 1990100680-A/1.		
SOURCE			
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 4378)		
TITLE	Ueda, K. and Komano, T.		
JOURNAL	HUMAN NORMAL CELL-DERIVED MDR RELATED GENE		
	SUNTORY LTD		
COMMENT	OS Homo sapiens		
	PN JP 1990100680-A/1		
	PD 12-APR-1990		
	PF 05-OCT-1988	JP 1988251475	
	PI UEDA KAZUMITSU, KOMANO TORU		
	PC C12N15/12, C12N1/21, C12Q1/68;		
	CC strandedness: Single;		
	CC topology: Linear;		
	CC *source: tissue=Adrenal gland;		
	CC *source: clone=SAM1132;		
	FH key	Location/Qualifiers	
	FT 5'UTR	1..137	
	FT CDS	138..3960	
	FT mat_peptide	138..3977	/gene='Multidrug resistance relating gene' FT
	FT	/gene='MDRA1'	
	FT	3'UTR	3981..>4378.
FEATURES	source	Location/Qualifiers	
		1..4378	
		/organism='Homo sapiens'	
		/db_xref='taxon:9606'	
BASE COUNT	1315 a	818 c	1057 g 1188 t
ORIGIN			
Query Match	36.3%	Score 1314.6;	DB 6; Length 4378;
Best Local Similarity	62.3%	Pred. No. 6-8e-266;	
Matches 2177;	Conservative	0; Mismatches 1269;	Indels 51; Gaps 5
QY	162	caagttgaccctgtatctatctgtggaatagtgctgctccttgatttctgttacaata	221
Db	473	CAGGTAGTGCCTATTATTACAGTGAATTTGGTGGCTGGCTGGTGGCTTACATTC	532
QY	222	gatttccttggtgattataactgcagcagcaagcaagcaagagattcgaaacagtttt	281
Db	533	GGTTTCATTTTGGTGCCTGGCAGCTGGGAACAATAATACAAATAATTAAGAAACAGTTT	592
QY	282	tcaatcagtttggcagacagacatcgctgctgttgatagctgtgacatcggtgaactta	341
Db	593	TCATGCTAATAATGCGACAGAGATATAGCGCTGGTTGATGTGCACGATGTTGGGAGCTTAA	652
QY	342	caatcgatgacagacatg--acaaaaatcagtatgatttgagataaagtctct	398

Db 653 CACCCACTTACAGATGATGCTCCAAAGATTATGAAAGGATGACAAAATTGGAAAT 712  
Oy 399 gttgttcaaaacatgtcaacttttcgatlttgccctgcaagtgttgtagaaggctg 458  
Db 713 gttctttagtcaatgacacatttttctactgggtttatagtagattttacactgttg 772  
Oy 459 gaactacacctagtgactctatccacgltctcctctataatgagctcagcggcaagatg 518  
Db 773 GAAGTAACCCCTGTGATTTTGGCATCAAGCTGCTTGTGAGACTGACCTGCTGCTG 832  
Oy 519 ttctaggaatggtcatctcatgaccgaatgaaggaattgaagcctatccaaactggg 578  
Db 833 GGCAGAAATCTATCTTCACTTACTGATAAAGAACTTTAGCGCTATGCCAAACCTGGAGC 892  
Oy 579 tgggagagaagaagctctgtcaccatccgaagatcatagacctttagggccgaaggaa 638  
Db 893 AGTAGCTGAAGAGGCTCTGGCAGCAATTAGAACTGTGATTTGGAGAGCAAAAGAA 952  
Oy 639 agaatctcaaaaglatacacagaatctcaagaatgcaaaagatttggcatalaaaggac 698  
Db 953 AGAAGCTTGAAGGTACACAAATAATTAGAAAGACTAAAGATGGATTAAGAAAGC 1012  
Oy 699 tataagctcaaaagtgctcttggtgctgtgtaactcttaatgaatggaactatggaact 758  
Db 1013 TATTACAGCCAAATATTCTATAGTGCTGCTTCTGCTGCTATCTATCTATCTATGCTCT 1072  
Oy 759 tgccttttgtagaagaccctcgtgacttcaatgaggaacccgtagatataccacggag 818  
Db 1073 GGGCTTTGGTATGAGGACCACTTGCTCTCAGGGGA-----ATAATCTATTGGACA 1126  
Oy 819 tgtctgtcgttcttctttagtgaatccatagagatlatgcatltgagacagacgtccc 878  
Db 1127 AGTACCTACAGTATTTCTTTCTGTATTAATTGGGGCTTTTAGTGTGGACAGCATCTCC 1186  
Oy 879 tcaacttgaacacttcgcaatagccggaagagctgcttccatacttccaggttatgta 938  
Db 1187 AAGGATTGAAGCATTTTCAATGCMAAGAGAGAGAGCTTATCAAAATCTTCAAGATAATTGA 1246  
Oy 939 taagaacccagatagatgaacttccacagctggagatatataccatgaatcacaagag 998  
Db 1247 TAATTAACCAAGTATTTCAGACCTATTTCGAAGAGTGGCACAACACATATTAATTAAGG 1306  
Oy 999 aactgtggaatttlaaaatgttcttcaatltacatccaagaccatctatcaaatctc 1058  
Db 1307 AAATTTGGAATTCAGAAATGTTCACTTACATTACCCATCTCGAAAAGAAAGTTAAGTCTT 1366  
Oy 1059 gaaagctgaatctcagaatlaagctcggagagaacagtcgcttggtcccaatgag 1118  
Db 1367 GAAGGGTCTGAACCTGAAGGTGACAGAGTGGCAGACGGTGGCTTGGTGAAGACAGTGG 1426  
Oy 1119 cagtggaagaagtgatcgttaagctcagcttctcagaagttatataatccggatgtgctt 1178  
Db 1427 CTGTGGGAAGGCAACAGTCCAGCTGATGCAGAGGCTCTATGACCCACACAGAGGGGAT 1486  
Oy 1179 tataatgtgagatagaatgaatcacaagagcttlaaatgtcgagcatlctcagagccat 1238  
Db 1487 GGTGAGTGTGATGGACAGATATTAGACATTAATGTTAGAGTTCATCAGGGAATCAT 1546  
Oy 1239 tggagtggttagtcaagaagcctgttctgctggagacacatcaagtaaatatcaagta 1298  
Db 1547 TGGTGTGTGATGACAGAACTGTATTGTTCACACGATAGTGAAGAACTATCGCTA 1606  
Oy 1299 tggagcagagatgagtgtcagtgtagaagatgagagagcaagaaggagaaatggtta 1358  
Db 1607 TGGCGGTGAAAATGTCCACATGATGAGATGGAAGAAAGCTCAAGGAAGCAATGGCTTA 1666  
Oy 1359 tgatttcaatcagagttctcctaataatlttaatacatgtgtaaggagaaagagctca 1418  
Db 1667 TGACTTTATCATGAACCTGCTCATTAATTTGACACCTGTTGGAGAGAGGGGCCCA 1726  
Oy 1419 aatgagtggaaggagaaacagagatcgcgaattgtcgtgcttagttcgaaccccaa 1478  
|||||

Db 1727 GTTGAGTGTGGGACAGACAGAGATCCGCATTGACGCGGCCCTGGTTGTAACCCAA 1786  
Oy 1479 gattctgattttagatgagagctacgtctccctgattcaagaagcaagtcaagcttca 1538  
Db 1787 GATCTCTCTGCTGGATGAGGCCACAGTACGCTTGGACACGAAAGCAAGCACTGGTTCA 1846  
Oy 1539 agctgcacgtggaagaaggcgagcaaaagctcgactaataatcggtgtagacaacgtcttc 1598  
Db 1847 GGTGGCTGGATGAGGCCGGAAGAGGTCGACCCCATTTGTATAGCTCATGCTTTGTC 1906  
Oy 1599 tactatcgaagtgcagatttgaatltgacccctaaagatggaatgctgctggagaaag 1658  
Db 1907 TACAGTCTGTAATGCTGACCTCATCGCTGTTGATGATGAGGATCATTTGTGAGAAAG 1966  
Oy 1659 agcacatgctgaactaatgcaaaacgaggtctatatattcactctgtgtaatgcaagga 1718  
Db 1967 AAATCATGATGAACCTCTGAAGAAAGAAAGCAATTAATCTTCAAACTGTCTCAATGACAG 2026  
Oy 1719 tattaataaagctgataagaatgagatgagatgaactatctca----- 1762  
Db 2027 AGCAGGAAATGAAAGTTGAATAATGACAGCTGATGATCAAAAGTAAATTGATGC 2086  
Oy 1763 -ctgaaagaagacaactcactctcctcgtactctgtgaagacatca----- 1810  
Db 2087 CTTGGAATGCTCTTCAATGATTTCAAGATCCAGCTATTAAGAAAAGATCAACTCGTAG 2146  
Oy 1811 -----agtcagacttcatgcaaaagctgaggaatccaaacacataagagat 1859  
Db 2147 GAGTGTCCGTGATCACAAAGCCCAAGAAAGCTTACTGACAAAGAGCTCTGATGA 2206  
Oy 1860 aagcttccctgaagctctctctatataaattttaaagttaaagaagcctgaatgagcctt 1919  
Db 2207 AAGTATACCTCCAGTCTTCTTGGAGATTGAACTAAATTAACTGAATGAGCCTTA 2266  
Oy 1920 tgtgtctgaggagacatgctcctcgtctcaaatgagacggttcatccatlatlccat 1979  
Db 2267 TTTTGTTGTTGTTGTTATTTTGTGCATTATTAATGAGAGGCTGCACACAGATTTGCAAT 2326  
Oy 1980 catcttgcgaanaatatcaacatgctt---ggaataatgatataaacacataaagca 2036  
Db 2327 AATATTTTCAAAAGATTAATGAGGCTTTTACAAAGAAATGATGATCTTGAACAAACGACA 2386  
Oy 2037 tgaatgcagaatattatccatgatattcgtcatcttgggtgtatgttctgttcaggtta 2096  
Db 2387 GAATGATTAATCTGTTTCACTATTGTTTCTAGCCCTTGGAAATTTTCTTTTATTAATTT 2446  
Oy 2097 ttcatcgaagattatttcaacgcagagcagaggaagaaatttcaacgaatgagataagaca 2156  
Db 2447 TTTCTCTTACAGGTTTACATTTGGCAAAAGCTGAGAGATCTTCACCAAGCGCTCCGATA 2506  
Oy 2157 ctgtgccttcaaacatgltatatacaggaatattgctgtgttgaatgaagaagaacag 2216  
Db 2507 CATGGTTTCCGATTCAGATGCTCAAGACAGATGATGATGTTGATGACCCCTTAAGAACAC 2566  
Oy 2217 caaagagagcttgcacaataatataagccatagatatagcaaaattcaaggaagaacag 2276  
Db 2567 CACTGGACATTTGACTACAGGCTCGCCAAATGATGCTCTAAAGTTAAAGGGCTATAGG 2626  
Oy 2277 ttcaagagatggtccttcaacaaatgcacaaatgcacaaatgcacaaatgcacaaatgcac 2336  
Db 2627 TTCCAGGCTTCTGTGAATTAATCCAGAAATATAGCAAAATTTGGACAGGAATTAATATTC 2686  
Oy 2337 ctlatatagatgagagatgacatltccgtatctcgtatgagatgctcagactgtccgt 2396  
Db 2687 CTTCATCTATGTTGGCAATTAACACTGTACTCTTACCAATTTGACCATCATATGGCAAT 2746  
Oy 2397 gacaggaatgattgaaccgagcaatgcatgatttgcacaaagaagtaagcaagact 2456  
Db 2747 AGCAGAGCTTGTGTAAGAAATGTTGTCGACAAACCTGAAAGATTAAGAAAGACT 2806  
Oy 2457 taagcatgctgaaagatagcaactggaagcttggagataatagctactatagtgtaact 2516  
Db 2807 AGAAGTCTTGGGAAGATCGCTACTGAAGCAATGAAAACTTCCGACACGTTGTTCTTT 2866  
|||||

OY	2517	aacaagggaanaaaccttcogagcaaalgtatgaagaagatgcttcagactcaacaagaa	2576
Db	2867	gactcagagacagaaagttttgaacatgtatgctcagaaatttgacaggtaccatcagaaa	2926
OY	2577	tactcgaagaagaacaaagatatttggaaagcgttatgatcatcgacatgctcttata	2636
Db	2927	ctcttttagaagaaacacacacatttcttggaaatpccattttccacccagggcaatgacta	2986
OY	2637	ttctgcctatgcagcaaggtctcgaatttggaccttaattcaagcttgcagcaatgac	2696
Db	2987	ttttttctatgctggaagttttccggtttggagccttactttggggcacaataactcatag	3046
OY	2697	cccaagaggacgltcaatagtttctacgtcaatttgcataltggagctatgcccacogaa	2756
Db	3047	cttttagagagatgttctgttatatttttccactgtttgctttggtgcattagccgtgggca	3106
OY	2757	aagctcgttttgccttcgttaattcccaagccaaltccggagctgcgtcatctgtttgc	2816
Db	3107	agtcagttcattttgctctgcacatctccaaagccaaaatpccagagccacatcatcat	3166
OY	2817	ctgttggaaaagaacacaatalatagacgcgcagtcacaaaggggaanaaagcaagac	2876
Db	3167	gatcatttgaaaaaaaccccttttgatttgacacgctracacacagaaagccctaatgccaacac	3226
OY	2877	altgtaagggaatttagagtltcgagaagtcctcttctctatccatgctgcgccaagat	2936
Db	3227	atttgaagaaatgtccattttggtgaattgtattcaactpccaccacccagcgacat	3286
OY	2937	tttaactctcogtgcgttacccttaagtatlttagcgagaagaagacagtagcatlttgg	2996
Db	3287	cccaagctcttgcagagactgcagccttggagggaagagccacagcgtgcgtctgtgtgg	3346
OY	2997	gagcagcgcgctgttggaaaagcacctcttcaactctgcagagacttatgacctgt	3056
Db	3347	cagcagtgctgcttgggaagacacacactgtgtccagctccttgagccggtttctacagccctt	3406
OY	3057	gcaaggacaagtgcgtgttlttagatgtgtgaatgcanaaagaattgaatgtacagtgtctcg	3116
Db	3407	ggcagagaaatgtgctgttatgttgacgaagaatpaaagccatgaaatgttcagtggtccg	3466
OY	3117	ttcccaaatagcaatcglttccctcaagaagcgtgtgctctcaactgcagcatgtcgtaga	3176
Db	3467	agcacacctgggcacatctctgtccacagagcccatctctgtttgactccacattgtcgana	3526
OY	3177	catgcgctatgtgtgaacaaagccgtgtgtgtcattagtagtgatcaagaagccgcgaa	3236
Db	3527	catttgcttatgtgaacacacagccgggtggtgtacaggaagaaatgttgagggcagcaaa	3586
OY	3237	tgcaagcaatatcatctcttltatgaaggtctcccttggaaatacaacaacaagaatttg	3296
Db	3587	ggagggccacatcatatgctctttcatgcaggtcacgtgcttaataatpatacagctaaagt	3646
OY	3297	actgaaggagcaacagctcttcctgcgcgcagaanaaacaagaactagctatgcgaagctc	3356
Db	3647	agacaaagaaactacgctctctgtgtggccagaaacaaacatttgcattagctgtgcctt	3706
OY	3357	ttctcaaaaaccanaatttatatttltgttatgtaggcaacttaagccctcgataatgaag	3416
Db	3707	tgtttagacagctcatattttttgcttttttgatgaagccacagctacgcttgatcagaana	3766
OY	3417	tgagaagtgltgttcagcatgccccttgaataagccaggaacggaagaagacatgctagtgt	3476
Db	3767	tgaaaaaggtttgtccaaagaccccttggaacaaagccagaaagccgcacotgatgtgat	3826
OY	3477	cactcaacagctctctgcgaattcaagaacgagatttgatagtgttctgcacaatytgaa	3536
Db	3827	tgcctacccgctctgcccacatccagatgcagacttmaatagtggttttcagaatgcag	3886
OY	3537	gataaagaanaaagaactataaagaagccctgcgagaatctgaagcatatatttlaagt	3596
Db	3887	agctaaagacatggcacgcatagcagctgctggcacgaanaagcactctatttttcaat	3946

QY	3597	agtgaaatgcacgctcag	3613
Dd	3947	GGTCAGTGTCCAGGCTG	3963
RESULT_14			
LOCUS	AX012321	8630 bp	DNA
DEFINITION	Sequence 2 from Patent EP0955374.		
ACCESSION	AX012321		PAT
VERSION	AX012321.1	GI:9998370	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	artificial sequence.		
REFERENCE	1 (bases 1 to 8630)		
AUTHORS	Baum,C.D., Hildinger,M. and Osterlag,W.P.		
TITLE	Retroviral vectors for gene transfer		
JOURNAL	Patent: EP 0955374-A 2 10-NOV-1999;		
FEATURES	HEINRICH PETTE INST (DE)		
source	Location/Qualifiers		
	1..8630	/organism="synthetic construct"	
		/db_xref="taxon:32630"	
		/note="Proviral Plasmid-DNA"	
	1..160		
misc_feature	/note="Plasmid-R ckgat (PUC)"		
misc_feature	1..8630		
	/note="Retroviralrler Expressionsvektor Sfpeta9lmsal"		
misc_feature	161..677		
	/note="5'-LTR"		
5'UTR	532..1219		
mat_peptide	1220..5062		
	/note="MSAI mdr1 cDNA"		
misc_feature	5215..5774		
	/note="3'-LTR"		
misc_feature	5775..8630		
	/note="Plasmid-R ckgat (PUC)"		
BASE COUNT	2234 a	2011 c	2189 g
ORIGIN	2196 t		
Query Match 36.3%; Score 1314.6; DB 6; Length 8630;			
Best Local Similarity 62.3%; Pred. NO. 6.8e-266;			
Matches 2177; Conservative 0; Mismatches 1269; Indels 51; Gaps 5,			
QY	162	caggttgcacctgtattatgtttggtaataaggctgttcgtccgttatlttttgtttacaataca	221
Dd	1555	CAGGTATGCCCTATTATAAGTGGAATTTGGTCTGGGGCGTGGTGCCTTACATTCA	1614
QY	222	gatcttcctgttgatatlaactgcagcacagacagaccagaagagattcgaaaacagtlttt	281
Dd	1615	GATTTCATTTTTGGTGGCTTGGCAGCTTGGAACAATAATACAAATAATTGAAAAACAGTTT	1674
QY	282	tcaatcagtttgcgcagcagagacatcgctcgtttcatagaatcgttgacatcgttgaaacttaa	341
Dd	1675	TCATGCTAATAATGCGCACAGAGAATAGGCTGGTTTGATGTGCACAGATGTTGGGAGACTTAA	1734
QY	342	cactcgcagtgcacagacatgtgac---aaaatcagtgatgtgatttgagaagaatgtgctc	398
Dd	1735	CACCCGACCTAACAGAGAGTGCTCTTAAGATTATGAAGTTATTGGGACAAAATTGGAAAT	1794
QY	399	gttgtttcaaaaabcttactcttttcgaatgagccgtgcagcttgatttggtgaagggtc	458
Dd	1795	GTTCTTTCAGTCAATGGCAACATTTTTCACCTGGGTTTATAAGTAGGATTTACACGTGGTTG	1854
QY	459	gaaacaccactagtgactctaatccaacyctcctcttataatgattgcagcggcagcatg	518
Dd	1855	GAAAGCTAACCCCTTGATGATTTTGGCCATCATGTCCTGTTCTTTGGACGTGACGTGCTGTG	1914
QY	519	tctcagtagtgatcatcactatgaccagtaagaatlaagtgcctattccaaagctggagc	578
Dd	1915	GGCAAAGTACATATCTTCATTTACGTATTAAGAAACCTTACGCTATGCAAAAGCTGGAGC	1974

OY	579	tgtggcagaagatctgcatcaaccgcacagtcataagctttagggcccaaggaa	638
Db	1975	AGTAGCTGAAGAGGTGTTGGCAGCATATTCGAAACCTGTGATTGCATTTGGAGCAAAAAGAA	2034
OY	639	agaactccaaggtatcacagatctcaaagtccaagatttgcatcaaaaagac	698
Db	2035	AGAACCTGAAAGGTACACAAAAAATTAGACAACCTTAAAGAATTGGGATTAAGAAAGC	2094
OY	699	tatatgtccaaaagtctctctggtgctgtaaccttttatgaattgnaacttggact	758
Db	2095	TATTACAGCCAAATATTCTATAGTGCTGCTTCCGTGGATCATATGATCTATTGCTCT	2154
OY	759	tgccttttgtatggaacctccttgacttcttaatygaaacctgatatcacatcgagac	818
Db	2155	GGCCCTTGCTATGGGACCCTTGGTCCCTCCAGGGG-----ATAATCTATTGGACA	2208
OY	819	tgtcttgctgcttttcttagtgtaatccatbagagattgatgcatitgtagagcaagccc	878
Db	2209	AGTACTACGTATCTCTTCTCTGTATTAATFGGGCTTTATGTGTGGACAGGACCTCC	2268
OY	879	tcacttgaaccttcgcaatgccgagagagctgcctttcatactttccaggttatga	938
Db	2269	AAGCATTTGAAGCATTTGCCAATGCAGAGAGACGCTTATGAAATCTTCAGATTAATTGA	2328
OY	939	taagaaacccagiatagataactcttccacagctgatalaaacctgaatccatagaag	998
Db	2329	TAAATAGCCAAAGTATGACAGTATTTCGAAGATGGGGCACAAACAGATATATTAAAGG	2388
OY	999	aactgtggaattaaaaaatgttctcttcaatatcatcaagaacatctatcaagattct	1058
Db	2389	AAATTTGGAAATCGAAATTTACTTCAGTTAACCCATCTCGAAAAGAAAGTTAAGATCTT	2448
OY	1059	gaaagtctgaatctcgaattagctcgtgagagacagctgcgcttgcttcgaatg	1118
Db	2449	GAAAGGCTGAACCTGAAGGTGCAGAGTGGGCAGAGCGTGGCCCTGTTGGAAACAGTGG	2508
OY	1119	cagtgaggaaagatcgygtagtcagcttctgcagaggtatatgaatccgatatbgctt	1178
Db	2509	CTGTGGAGAGGCACACACACTCCAGCTGATGCGAGAGGCTCTATTGACCCCAAGAGGGGAT	2568
OY	1179	tatatgltgatbatagatatbaotcogagctttaaatltyggsgaatlatlgagaccatat	1238
Db	2569	GGTCACTGTTATGTAGACAGGATTTTGGACCATTAATGTAAAGTTTCTAACGGGAATCAT	2628
OY	1239	tggaigtgttagtcaaaagagctgttctgttcygaaccacacaglaaacatatcaagta	1298
Db	2629	TGCTGTGCTGATGAGGAACTGATTTGTTGGCACACAGTAAAGCTGAAACATTCGCTA	2688
OY	1299	tggaagagatgatgtgcactgaagaagatbtgagagaagcagcaagggaaagacaatgtcta	1358
Db	2689	TGGCCGTGAAATATGTCAACATGTGATGATTTGGAAGAGCTGTCCAAGGAAGCCAAATGCCTA	2748
OY	1359	tgattttatcatgtagtcttccatbaatttaaacaattgfrtagggaaaaagggactca	1418
Db	2749	TGACTTTATCAVGAACCTGCTCATTAATTTGACACCTTGTTGGAGAGAGAGGGGCCCA	2808
OY	1419	aatggttggaaggaagaacaagagatcgaatctgtctgttgcttaagtgcgaaccocaa	1478
Db	2809	GTTTAGTGTGGTGGCGAAGACAGAGAGATCGCCATTTGACAGCTGCCCTGTTGCCAACCCAA	2868
OY	1479	gattctgattttagatgaggtactgctcctgtgatcagaagaagtcagctgttca	1538
Db	2869	GATCTCTCTGCTGGATGAGGCCAGCTCAGCCTTGACACAGAAAGCGAAGCAGTGGTTCA	2928
OY	1539	agctgactatggaagagcgagcaaaagttcgactacaatcgtggttagcaacacagcttcc	1598
Db	2929	GGTGACTCTGTGATAGGCCA AAAAGGTGCGACACCATTTGATATAGCTCATCTGTTGTC	2988
OY	1599	tactcttgaagbtcaattgattgtgaccccaaaagatgagatctgcygagagaag	1658
Db	2989	TACACTTGTGTAATGCTGACGTCATCGCTGGTTTTCAATGATGAGGATCATCTGTGGAGAAAGG	3048

QY	1659	agccatcgtgaactaatgycacaacgaggtcatatattcaactctgtgactggtcacaagga	1718
Db	3049	AAATCATGATGAACTCATGAAAGAGAAAGGCATTTTACTTCAAACTGTGCATAATGCAGAC	3108
QY	1719	tattaaaagcgtgatbaacaagatgtagcacaatgcatactcta-----	1762
Db	3109	AGCAGAAATGAATTGAATTAGAAAATGCAGCTGATGATACCAAAAGTGAAATTGATGC	3168
QY	1763	-ctgaagaagaagaactaactactccctcgtcaactctgtgaagagcatca-----	1810
Db	3169	CTTGGAATAATGCTTCAAAATGATTCACAGATTCAGAGTCTAATTAAGAAAATGATCACTCGTAG	3228
QY	1811	-----agtcagaactcattgacaagagctgaagaaatccaccacttaagaagat	1859
Db	3229	GAGTGTCGTGATCACACAAACCCAAACAGAAAGCTTTAGTACCAAGAGGCTGTGATGA	3288
QY	1860	aagcttcctgaagtcctctctattaaaatttaagttaaacaagcctgaatgacctt	1919
Db	3289	AAGTATACCTCCACTTTCCTTTTGGAGATTTAGAGCTAAATTTTACTGAAATGGCTTA	3348
QY	1920	tgtagttctgaggagactcgtgctctgcttcaaatggaactgtctccagatattccat	1979
Db	3349	TTTTGTCTTGTGGTATTTTGTGCTCAATTAATATGAGGCCCTGCACACAGCATTTGCAAT	3408
QY	1980	catcttgcacaaaatltataaccatgltt--ggaataatgataaaccacattaaagca	2036
Db	3409	AATATTTCAAAAGATTTATAGGGGTTTTACAAAGATTTGATGATCCTGCAAAAACGACA	3468
QY	2037	tgatgcagaatattatccatgatatactgctcaatttgggtgtatttgccttgccagta	2096
Db	3469	GAAATGTAACCTGTTTTACATATGTTTTTCAGCCCTTGAATTAATTTCTTTATTTACATT	3528
QY	2097	ttctcgaagagattatttcaagcagagcagaggggaatcttcaagatgagattaaagca	2156
Db	3529	TTTTCTTAAAGCTTTCACATTTTGGCAAGCTGAGAGATCCCTACCACAGCGCTCCGATA	3588
QY	2157	cttgaccttcaaacccatgltatatacgaatatgctcgttggttgaatgtaaaaggaacag	2216
Db	3589	CATGGTTTTCCGATTCATAGCTACAGACAGATGAGATTTGGTTGATGACCCATAAAACAC	3648
QY	2217	cacagagagcttgacacaatatattagccatagatataagacacaattcaagagcacaag	2276
Db	3649	CAGTGGACATTGACACCCAGGCTGCGCAATGATGCTGTCAAGTTAAAGGGGCTATAGG	3708
QY	2277	ttccagagattgagcgtcttaacacaanaatgacaatacaatcagggacttcagttatcattc	2336
Db	3709	TTTCCAGGCTTCTGTATNTATNCCGAGATATAGAGAAATCTGGACACGAAATATTATATTC	3768
QY	2337	ctttatataatgatatggaagatgacatctcgtattctgagatattgctccagtaactgccc	2396
Db	3769	CTTCATCTATGTTGGCACTAACACACTGTACTTACCAATTTGTAACCCATCATTTGCAAT	3828
QY	2397	gacaggaatgattgaacccgagaacatgactgtagatttgcacaacaagaataagcaagact	2456
Db	3829	AGCAGAGAGTGTGTAATGAAGAAATGTTGTCTGGACAAGCACTGAAAGATTAAGAAAGACT	3888
QY	2457	taagatctctgaagaagatgacaactctggaagcttgggaaatatacgtactatagtgcat	2516
Db	3889	AAGAGTGCTGGGAAGATCGCTACTGAGCAATATAGAAAACCTTCCAGACCGTGTTCCTT	3948
QY	2517	aacaaggggaaaaagccttcgagcaaatgtatagagaagatgctcagaactcaacagagaa	2576
Db	3949	GACTCAGGAGCGAAGTTTGAAACTATGTATGCTCAGAGCTTGACAGGTACCATACAGAAA	4008
QY	2577	tacctcgaagaagacacagatatttggaaagctgttatgcatcttcagcatcagccttata	2636
Db	4009	CTCTTTGAGGAAGCACACATCTTTGGAATTTACATTTTCCCTTCCACCGACATATATGTA	4068
QY	2637	ttttccctatgaagcaggggttctgatttggagccttattcaatcaagcctggaagcatgac	2696
Db	4069	TTTTTTCCTATGCTGAGATGTTTTCGGGTTTGGAGCCTACTTGTGTGGCAATAAATCATGAG	4128
QY	2697	cccaagaggcatgltcatagatttttctacgcaattgcatatggaagctatgagccatcggaa	2756

Db	4129	CTTTGAGGATGTTCTGTTAGTATTTTTCAGCTGTTGTCCTTTGGTCATGCGCTGGGCGA	188
Oy	2757	aagcctcglttttgctctctgaatcattccaaagccaatcggggcctgcgcattcttgc	2816
Db	4189	AGTCAGATTGATTTCCTCTCTACATATGCGAAAGCCAAATATACGACGCCCATCATCATC	4248
Oy	2817	cttgttggaaaagaaccaaataatagacgcgcagtcgaagaaggaaaacacagac	2876
Db	4249	GATCATTTGAAAAAACCCTTTGATTGATTCAGACTACACGACGGAAGGCGTTAATCCGAAAC	4308
Oy	2877	atggaagggaatttagagtttcgagaagctctctctctctcattcgaatgcccgaagt	2936
Db	4309	ATTGGAAAGGAATCTCATTCTGGTGAAGTTGATTCAACTATCCACCCTGACGGACAT	4368
Oy	2937	tttcatcctccgtygtcttaaccctcgaattgagcaggaaaagacagtagcaatttgg	2996
Db	4369	CCCAAGTGCTTCAGGAGACTGAGCCTGGAGGTGAAGAAAGGCCACAGCGTCTGTGTGG	4428
Oy	2997	gagcagcgcgtctgaggaaaagacctctgttcaactcttcgagagacttattgacccgt	3056
Db	4429	CAGCAGTGCGCTGTGGGAAGGACACAGTGGTCCAGCTCTCTGAGCGGTTTACGACCCTT	4488
Oy	3057	gcaaggagaagttgctgtttagttgttgatgtaaaagaattgtaatgtacagtgtctcg	3116
Db	4489	GGCAGGGAAGATGCTGCTTATGTCGCAAAAGAAATTAAGCCAGCTGAATGTTCACTGGCTCG	4548
Oy	3117	ttcccaaatagcaatctgtcttccaagagccctgctgtccttcaactgagatctgcgagaa	3176
Db	4549	AGCAGACCTTGCGCATCTGTCGCCAGAGCCCATCTGTTTGACTCGAGATTGCTGAGAA	4608
Oy	3177	catgcctatgtgtgacaacagccgtygtgtyccattagatgtagatcgaagaagccgcgcaa	3236
Db	4609	CATTGCTATGAGAGACAAACAGCCGGGGTGTCTACAGGAAGAGATCGTAGGGCAGCAAA	4668
Oy	3237	tgcgcgaatatcatctcttatttgaaggtctccctcgagaatatcaacacaaagttgg	3296
Db	4699	GGAGGCCAACTATACATCCCTTCATCGAGTACCTGCTTAATAATATAGCATTAAGTAGG	4728
Oy	3297	actgaagaaggagcagctcttctgcgcgccaagaacaaagactagctattgcaaggcct	3356
Db	4729	AGACAAAGAGACTACGTCTCTGTGTGGCCGAAACAACCATTTGCCATAGCTGTGGCCT	4788
Oy	3357	tctccaaaaacccaaaatttatttgttgaatgtaggcaccttaagccctcgataatgacag	3416
Db	4789	TGTTAGACAGCCTCATATTTTGGCTTTGGATGAAGCCACGTCAGCTCTGGATACGAAGA	4848
Oy	3417	tggagaagttgttttagatgatgccccttgataaagccaggaaggaagacatgctcgtagt	3476
Db	4849	TGAAAAGGTTGTCACAAACACCTCTGCAAAAGGCAAGAGGCCGCCGACCTGCACTGTAT	4908
Oy	3477	caatcacaaggtctctctgcaattcaagaagcagatttggatagttgtctgcacatggaaa	3536
Db	4909	TGCTCACCGCCTGTCCACACTCCAGATGCAAGTGAAGACTTAATATGTGTGTTTCAATAATGGCAG	4966
Oy	3537	gataaaggaaacaaagcaatcacaagaagctcctgagaanaatcgagacataatttgaatt	3596
Db	4969	AGTCAAGAGCATGGCAGCGATCAGAGCTGTGTGCACAGAAAGGATCTATTTTTCAT	5026
Oy	3597	agtgaatgcaaatgcag	3613
Db	5029	GGTCAGTGTCCAGGCTG	5045
RESULT 15			
AX012320	AX012320	8630 bp	DNA
DEFINITION	Sequence 1 from Patent EP0955374.	PAT	06-SEP-2000
ACCESSION	AX012320		
VERSION	AX012320.1	GI:9998369	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		

REFERENCE	1 (bases 1 to 8630)	artificial sequence.
AUTHORS	Baum,C.D., Hildinger,M. and Osterlag,W.P.	
TITLE	Retrieval vectors for gene transfer	
JOURNAL	Patent: EP 0955374-A 1 10-NOV-1999;	
	HEINRICH PETTE INST. (DE)	
FEATURES		
source	1..8630	Location/Qualifiers
	/organism="synthetic construct"	
	/db_xref="taxon:32630"	
	/note="Proviral Plasmid-DNA"	
misc_feature	1..8630	
	/note="retroviral/er Expressionsvektor Sfbeta71m4"	
misc_feature	1..160	
	/note="Plasmid-R ckgat (puc)"	
misc_feature	161..677	
	/note="5'-LTR"	
5'UTR	532..1219	
mat_peptide	1220..5062	
	/note="m4 mdr-1 cDNA"	
misc_feature	5215..5774	
	/note="3'-LTR"	
misc_feature	5775..8630	
	/note="Plasmid-R ckgat (puc)"	
BASE COUNT	2235 a 2011 c 2187 g 2197 t	
ORIGIN		
Query Match	36.2%; Score 1311.4; DB 6; Length 8630;	
Best Local Similarity	62.2%; Pred. NO. 3.2e-265;	
Matches 2175; Conservative	0; Mismatches 1271; Indels 51; Gaps	5;
QY	162 caggttaccgcggtatcttattgttgatagatggttgctgccttgatcttttggttacaata	221
DB	1555 CAGGATATGGCTATTATTATAGTGGAAATTGCTGCGGCGCTGCTTACATTCA	1614
QY	222 gattcccttgatgatatataactctgcacagacagacccaagaggatctcgaaaacagtttc	281
DB	1615 GGTTCATTTTGGTCCCTGGCAGCTGGAAGACAAATTACAAATAATTAGAAACAGTTT	1674
QY	282 tcatccagtttggcacaggacatcgcctggttgatgctgtgtgacatcgcgtagaacttaa	341
DB	1675 TCATGCTATTAATGCGACAGAGATAGCGTGGTTGATGTGACGAGATGTGGGAGCTTAA	1734
QY	342 caactgcacatgacacatctgac---aaatcagtgatggttatgtgataagatctct	398
DB	1735 CACCGGACTTACAGATGATGATGCTCTAAGATTATAGATTATTTGGTGAACAAATTGGAA	1794
QY	399 gtcttctcaaaaacatgctactcttctcgattgctgcctgcagcttgctgtgtaaggctg	458
DB	1795 GTTCTTTTCAGTAAAGGCAACATTTTTCACCTGGTTTATAGAGGATTTACACGTGGT	1854
QY	459 gaaactacccttaagtgactctatccacgctctctcttcttlaatgacttgaagcgagcatg	518
DB	1855 GAAGCTAACCCCTTGATTTTGGCCATCAGTCTGTTCTTGAGCTGCAGCTGCTGCTG	1914
QY	519 tctcaggaatgcatcatctatgccacgaatgaatgaatgaatgaatgaatgaatgaatgaat	578
DB	1915 GCGCAAGATACATCTCTCTCTTACCTAGATTAACAACTTTAGCGTATGCAAAAGCTGGAGC	1974
QY	579 tctggcagaagaagctctgtgtcattcaatcccgaaacagtcataagcctttagggcccaagaa	638
DB	1975 AGTAGCTGAAGGAGGCTTGGCAGCAATTAGACCTGATTCATTTGAGAGCAAAAGAA	2034
QY	639 aagaactcaaaagtatacagaatctcaaaatgcaaaagatcttggcacaataaaagac	698
DB	2035 AAGAACTTGAAGAGTACAAACAAATTTTGAACAAACTTAAAGAAATTTGGATTAAGAAAGC	2094
QY	699 tatagcttcaaaagtgctctgt	758
DB	2095 TATTACAGCAATTTTCTATAGTGCTGCTTCTTCTGCTGATCTATGCATCTTATAGCTCT	2154
QY	759 tgccttttgatagtgaacctcttgattcttaatgtgagaaccttgatataacatccgggac	818

Db	2155	GGCCTTCTGTAATGGGACACACCTTGGTCTCTTCAGGGGA-----ATAATCTATTGGACA	2208
Qy	819	tgcttcgtcgtcttccttcttagtgtaatccatagcagatlatcgcattcgggcagcagctcc	878
Db	2209	AGTACTACTGATATCTTTCTTCGTATTAATATGGGGCTTTAGTGTTGACAGCACTCTCC	2268
Qy	879	tcaacttgaacacctgcgaatagcccgagagctgccttcatatttccaggtatga	938
Db	2269	AAGCATTTGAGCATTTGGCAAAATGCAGAGAGCAGCTTATGAAATCTTCAAGATTAATTGA	2328
Qy	939	taagaaccccgatagataacttttccacagctgataataaccgtaaccgaatgaagaag	998
Db	2329	TAATAGCCCAATATTGACAGCTATTTCGAAGAGTGGGCACAAACCAGATATATTTAAGG	2388
Qy	999	aactctggaatttaaaatggttcttcttcaattatccatcaagacacctcaagaattc	1058
Db	2389	AAATTTGGAAATTCAGAAATGTTCACTTCCAGTTACCATCTCGAAAGAAGTTAAGATCTT	2448
Qy	1059	gaaagctctgaactcgaataaagtaagtcgtagagagacagctgcgcttgcgttcgaatg	1118
Db	2449	GAAGGCGCTGAACCTGAAAGGTGCAGAGTGGGCAGAGCTGTATGAGCCCTCGTTGGAACAGTGG	2508
Qy	1119	cagtgaggagatgaogttagctcagcttctcgagaggtatataagaccgatatgagctt	1178
Db	2509	CTGTGGGAAGACACAAACGCTCAGCTGATGCAAGAGCTCTATGAGCCCAACAAGGGGAT	2568
Qy	1179	tatcatgtgtgataagaaatgacatcagagctttaaigtgcgcattatcgagacat	1238
Db	2569	GGTCACTGTTGATGAGCAGAGATATTGACACAAATGTAAGTTTCTACGGGAATTCAT	2628
Qy	1239	tggagtggttagtcaagagccgttcttgctcgggaccacatcagtaacaatatcaagta	1298
Db	2629	TGGTGTGGAGATCCAGAACCTGTATGTTGGCCACAGCATAGCTGAAACAATTCGCTA	2688
Qy	1299	tggagagagatgatlgyacatgaaagagatlgagagagcagcaaggaagcaaatggtta	1358
Db	2689	TGGCGGTGAAATGTCACATGGATGAGATTGAAAGCGTCAAGGAAGCCCAATGCTTA	2748
Qy	1359	tgaatttcaatcggagcttccctaataatttaaatacatgtgtaggggaaaaagagctca	1418
Db	2749	TGACTTTATCATGAAACCTGCCTCATTAATTTGACACCCCTGGTTGGACAGAGAGGGCCCA	2808
Qy	1419	aatgagtggaaggcagaaaaacagagatcgcgaattgcgtgccttagtttcgaaaccccaa	1478
Db	2809	GTTTGATGGTGAGACAAACAGAGATGCCATTGGACAGCGCCCTGTTGCAACCCCAA	2868
Qy	1479	gattctgatttagatgagagctacgctcgtccctcgattcgaagaagcaagtccagctgtca	1538
Db	2869	GATCCTCCCTCGTGGAGAGGCCACGTCAGCCTTGGACACGAAAGCAACACAGTGTTC	2928
Qy	1539	agctgcacggagaaagcgagcaaaagtcggaactacaatcgttgtagcacacgaattc	1598
Db	2929	GGTGGCTCTGGATAAGGCCAGAAAAGTGGACACCACCATTTGTATAGCTATGCTTTGTC	2988
Qy	1599	tactatctgaagtgcagatttgaatgtgaacctaaagatgaaatgctgcgggaaag	1658
Db	2989	TACAGTTCGTAAAGTGACGTATCGCTGCTTGATGATGAGATCATTTGTGGAGAAGG	3048
Qy	1659	agcacatcgtgaactaaatggcacaacgaggtctatatattcaatcctgtgtatgcagga	1718
Db	3049	AAATCATGTATGAACATCATGAAGAAAGGCAATTTACTTCAAACTGTCTCAATGCGAC	3108
Qy	1719	tattaaaaagcgtgagaaacagatgagatgaacatattca-----	1762
Db	3109	ACCCAGAAATGAAATTGAATTAAGAAAATGCAGCTGATGAATCCAAAAGTGAATGTATGC	3168
Qy	1763	-cttgaagaagaagcaactactcctctgcactcgttgaagagcatca-----	1810
Db	3169	CTTTGGAAATGTCTTCAAAATGATTTCAAGATCCAGTCTTAATTAAGAAAAAGATCAACTCGTAG	3228
Qy	1811	-----agtcgaactcatctgacaagagctlgagaaatcaccaccaatctaaagat	1859

Db	3229	GAGGTCCCGTGATTCACAAAGCCCAAGACAAAGACTTAGTACCAAAAGGCTCTGATGA	3288
Qy	1860	aagctcttcgaagctctctataltaaaatlttaagltlaacaagcctgaatgctctt	1919
Db	3289	AAGTATACCTCCAGTTTCCCTTTTGAGAGATTAATGAAGTAATTAATCTGAATGGCCCTTA	3348
Qy	1920	tgtgtcttcgggacatctgctctgtcttaatggaactgttcatccagatlttccat	1979
Db	3349	TTTTGTTGTTGGTATTTTGTGCAATTAATAAGAGGCTCGCAACCGCATTTTGCAAT	3408
Qy	1980	catctctgcaaaatltlaacacagttt---ggaatatgataaacaacatitaaagca	2036
Db	3409	AATATTTTCAAGATTTATAGGGGTTTTTACAGAAATGATATGATCTGTAAACAAACGACA	3466
Qy	2037	tgatgcagaatlatatccatgatatctgcgaatttgggtgtatctgtcttgcagtta	2096
Db	3469	GAATAGTAACCTTGTTCACATATTGTTCTTGAGCCCTTGGAATTAATTTCTTTATTAACAT	3528
Qy	2097	tttcatgcaggtatattttacgcgacagcgagggaattttaaactgagatlaagaca	2156
Db	3529	TTTTCTTCAAGTTTCACTTATTTGGCAAGCTGGAGAGATCCTCACACAGCGGCTCCGATA	3586
Qy	2157	cttgcccttcaaacagatgtatatacaggaatctgcgtgtttagtgaagaaagaaacaag	2216
Db	3589	CATGCTTTCCGATCCATCGTCACACAGAGATGATGATGGTTGATGATGCCCTAAACAC	3646
Qy	2217	cacagagagcttgaacaacatattagcatagatatagacacaatltcaagaagcaacag	2276
Db	3649	CACGTGAGCATTTTGACTACACAGCGCCCAAGATGCTCTCAAGTTAAAGGGCGCTATAGG	3706
Qy	2277	ttccaagatggcgtcttaacacaaatgcaactaacctggagacttcaagtatcatctc	2336
Db	3709	TTTCCAGCTTCCTTAAATTACCAGAAATATGCAAAATCTTGGACAGCAATTAATTAATTC	3766
Qy	2337	ctttatatgtgactggagagatcaatctcctatctcgtatctgcttgcagtaactgcgct	2396
Db	3769	CTTATCTATGTTGGCAGACTTAACACTTTACTCTTACCAATTTGATCCCATCATTTGCAAT	3826
Qy	2397	gacagagaatgtatgaacccgagcaatgactggaatttgcacaaagaataagcaagaact	2456
Db	3829	AGCAGAGATTGTGTAATGAATAATGTTGTGACACACCTGAAAGATTAAGAAAGAACT	3886
Qy	2457	taagcatgctggaaagataagcaactggaagcttggagaatatagttactatagtgctat	2516
Db	3889	AGAAAGTCTGTGGAAAGCTGTACTGGAAGCAATTAAGAACTTCCAAACCGTGTTCCTT	3946
Qy	2517	aacaaggaagaaagccctcgagcaaatgtatgaagagatgctcagacatcaacacagaa	2576
Db	3949	GACTCAGAGCAGAAAGTTTGACATATATATGCTCAAGATTTGGCAGATTCATACAGAA	4006
Qy	2577	tacctcgagaagacacagatlatgtgaagctgtatgacatcagcacatgccttatata	2636
Db	4009	CTCTTTTGAGAAAGCACATCTTTGGAATTAACATTTTCTTCACCCAGGCATGATGTA	4066
Qy	2637	tttgctcatgcagcaggttctgaatttggagccattatattcaagcttgcagatgac	2696
Db	4069	TTTTTCCATCTGAGATTTTCCGTTTGGAGCCCTATCTGTGGCAATAACTCATGAG	4126
Qy	2697	cccaagagg	





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:45:58 ; Search time 293.81 Seconds  
(without alignments)  
1557,850 Million cell updates/sec

Title: US-09-873-409-15

Perfect score: 2021  
Sequence: 1 ttccgcttgcgtatgact.....taaaatgctactttatga 2021

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/Backfiles1.seq:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	512.8	25.4	4646	1	US-08-181-471-2
2	512.8	25.4	4669	2	US-08-752-447-1
3	512.8	25.4	6505	2	US-08-793-610-5
4	512.8	25.4	9318	2	US-08-793-610-6
5	511.2	25.3	4669	6	5206352-3
6	506.4	25.1	4669	2	US-08-583-276-18
7	503.4	24.9	4264	2	US-08-784-649A-1
8	501.8	24.8	4264	2	US-08-784-649A-5
9	477.4	23.6	4233	3	US-09-120-513-1
10	477.4	23.6	4233	4	US-09-450-105-1
11	221.8	11.0	2726	1	US-08-461-823-1
12	175.2	8.7	4002	2	US-08-996-545-1
13	175.2	8.7	4002	2	US-08-996-545-3
14	175.2	8.7	4002	2	US-09-328-320-1
15	175.2	8.7	4002	4	US-09-328-320-3
16	170	8.4	4224	1	US-08-612-521-1
17	168.6	8.3	4047	2	US-08-612-734B-1
18	168.6	8.3	4800	2	US-08-612-734B-3
19	143.8	7.1	3924	1	US-08-395-246C-1
20	142.8	7.1	2061	4	US-09-061-764A-17
21	141.2	7.0	1959	4	US-09-061-764A-4
22	121.6	6.0	2376	1	US-08-394-880B-1
23	116	5.7	3924	2	US-08-996-644-3
24	116	5.7	3924	3	US-09-352-552-3
25	116	5.7	3927	2	US-08-996-644-1
26	116	5.7	3927	3	US-09-352-552-1
27	105.4	5.2	5120	3	US-08-772-270A-6

28	105.4	5.2	8370	2	US-08-488-706-1	Sequence 1, Appl
29	105.2	5.2	2244	4	US-09-061-764A-18	Sequence 18, Appl
30	104.4	5.2	4403765	4	US-09-103-840A-2	Sequence 2, Appl
31	104.4	5.2	4411529	4	US-09-103-840A-1	Sequence 1, Appl
32	103.6	5.1	6143	1	US-08-612-521-3	Sequence 3, Appl
33	100	4.9	7721	3	US-08-772-270A-14	Sequence 14, Appl
34	98.2	4.9	2407	3	US-08-895-522-2	Sequence 2, Appl
35	98.2	4.9	2407	3	US-09-195-391-2	Sequence 2, Appl
36	91	4.5	4403765	4	US-09-103-840A-2	Sequence 2, Appl
37	90.6	4.5	3909	1	US-08-232-537-1	Sequence 1, Appl
38	89.4	4.4	28804	2	US-08-592-874-1	Sequence 1, Appl
39	89.4	4.4	28804	3	US-09-096-942-2	Sequence 2, Appl
40	89.4	4.4	28804	3	US-09-096-867-2	Sequence 2, Appl
41	86.4	4.3	1201	3	US-08-961-083-189	Sequence 189, App
42	86	4.3	5889	1	US-08-463-092B-5	Sequence 5, Appl
43	86	4.3	5889	2	US-08-462-109A-5	Sequence 5, Appl
44	86	4.3	5889	2	US-08-460-907B-5	Sequence 5, Appl
45	86	4.3	5889	3	US-08-463-179A-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-08-181-471-2  
; Sequence 2, Application US/08181471  
; Patent No. 5641508  
; GENERAL INFORMATION:  
; APPLICANT: Li, Lingna  
; APPLICANT: Lishko, Valery K.  
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESSES:  
; ADDRESS: Thomas Fitting  
; STREET: 12526 High Bluff Drive, Suite 300  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92130  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/181,471  
; FILING DATE: 13-JAN-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/041,553  
; FILING DATE: 02-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: ANT0029P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-792-3680  
; TELEFAX: 619-792-8477  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4646 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 425..4267  
; US-08-181-471-2



FEATURE:  
NAME/KEY: 5'UMR  
LOCATION: 1..424  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4264  
FEATURE:  
NAME/KEY: 3'UMR  
LOCATION: 4265..4669  
US-08-752-447-1

Query Match 25.4%; Score 512.8; DB 2: Length 4669;  
Best Local Similarity 62.0%; Pred No. 3.5e-143;  
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;

QY 162 caggttgaacctgtatattgttgaaatagtgctgcctctgatttttggtaataca 221  
DB 760 CAGGTATGCGCTATTATTACAGTGAATGCTGCTGGGGTGTGCTTACATTCA 819  
QY 222 gatttcctgtgtatataactgcagacacagaccagaagatcgaacaagctttt 281  
DB 820 GATTTCATTTTGTGCTGGCAGCTGGAGACAAATACAAATTAAGAAACAGTTT 879  
QY 282 tcatcagtttggacagacatcgctggtttgaatagctgtgacatcggtgaactta 341  
DB 880 TCATGCTATTATAGCGACAGAGATAGGCTGTTGATGTGACAGATGTTGGAGCTTAA 939  
QY 342 cactgcacatgacagacatg---acaaatcagtgatggtatgtgaataagatgctct 398  
DB 940 CACCGGACTTACAGATGATGCTGCCAAGATTATGAAGAAATTTGGACAAATTTGGANT 999  
QY 399 gtgtttcaaaaacatgctcacttttgcattggtcctgagctgtgttggtaagggctg 458  
DB 1000 GTTCTTTCAGTCAATGGCAACATTTTTCACGTGGTTATAGTAGATTTTACAGGTGTTG 1059  
QY 459 gaaactcacccatgagactctacacgctcctcttataatgagctcagcgagacatg 518  
DB 1060 GAAAGCTAACCCCTTGATGATTTTGCCATCAGTCTGTTTGGACGTGTCAGCTGCTCTG 1119  
QY 519 ttctcagatgctcatctcatctacatgacaglaagaaatgaatgaatgctcattcaaa 578  
DB 1120 GCGAAGAGTACTATCTCTCATTTACGATTAAGAACTCTTAGCGTATGCAAAAGCTGGAGC 1179  
QY 579 tgtgacagaagaagctctgtcatcaatccgaacatgacatgacgtttagggccagaaga 638  
DB 1180 AGTACTGGAAGAGGCTTGGCAATTAAGAACTGATTTGATTTGGAGGACAAAGAA 1239  
QY 639 agaactccaagaagctcttctttaaataaagaatagctgtgtttatttcccca 698  
DB 1240 AGAAGTTGA----- 1248  
QY 699 gtggtactaagtgltgtlctgtlntltglaagatatacagaatctcaaatgataca 758  
DB 1249 -----AAGGTACACAAATAATTAGAAAGAACCTAA 1278  
QY 759 ggaatttggcataaaaagaactatagcttcaaaagtctctgtgtgtgtacttctt 818  
DB 1279 AAGAAATTGGGATTAAGAAAGCTATTACACCAATATTCTATAGGCGTCTTCTGCT 1338  
QY 819 tatgaatgaactatgagctgtcttggatgtaagaaactctcattcctaataatgaga 878  
DB 1338 GATCTATGCAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1398  
QY 879 accgtgatatacaccatcggaagctgtcttcttcttcttcttcttcttcttcttctt 938  
DB 1399 -----ATTCTTATGGAAGAAAGTACTCACTGATTTCTTTTCTGATTAATTTGGGCTTT 1452  
QY 939 ttgcatgtgagacagagctccctcacttgaacacttgcgaataagccgagagagctt 998  
DB 1453 TAGGTGTGACAGGCAATTCGCAAGCAATTTGCAAAATGCAAGAGGAGACACCTTA 1512  
QY 999 tcatatttccaggttatgataagaaccagtatagtaacttcttccaaagcttgata 1058

DB 1513 TGAATCTTCAAGATATATTATTAAGCCAACTATTGACAGCTATTCCAGAGTGCGCA 1572  
QY 1059 taaacctgaatccatagaagaagctgtgaaattaaatgtttcttcaatccatc 1118  
DB 1573 CAACACAGATTAATTAAGGGAATTTGGAAATTCAGAAATGTTACATTGACATTACCATC 1632  
QY 1119 aagaccatctataagaatcttgaagaagcttgaatctcagaatgaatgtctggaagacagt 1178  
DB 1633 TCGAAAAAGATTAAAGATCTTGAAGGCGCTGAACCTGAAGCTGACAGTGGGAGACGGT 1692  
QY 1179 cgcctgtgcagtcgaatgagcagtgagatgagatgagatgagatgagatgagatgagat 1238  
DB 1693 GCGCTGCTGTTGGAAGAGTGTGCTGAGAGACACAAAGTGTGCTGAGAGAGCT 1752  
QY 1239 atatgaccgagatgagcttcaatgctgtgagatgagatgagatgagatgagatgagat 1298  
DB 1753 CTATGACCCACAGAGGAGTGTGCTGAGTGTGATGAGACAGATATTAGACACATTAATGT 1812  
QY 1299 gggacattatcgagacatattgagatgagatgagatgagatgagatgagatgagatgagat 1358  
DB 1813 AAGGTTTCAAGGGAATCAATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1872  
QY 1359 catcagtaacatatacgaatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1418  
DB 1873 GATAGCTGAAGAAATTCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932  
QY 1419 agcaagaggaagcaaatgctgattatctatcagatgagatgagatgagatgagatgagatgagat 1478  
DB 1933 TGTCAAGGAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1992  
QY 1479 ggtaggggaaagagagctcgaatgagatgagatgagatgagatgagatgagatgagatgagat 1538  
DB 1993 GGTGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2052  
QY 1539 tgcctatgctcgaaccccaagatctcgaatgagatgagatgagatgagatgagatgagatgagat 1598  
DB 2053 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2112  
QY 1599 agaaagcaagctcagctgttcaagctgcaactggaagag 1636  
DB 2113 AGAAGCGAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2150

RESULT 3  
US-08-793-610-5  
Sequence 5, Application US/08793610  
Patent No. 5858744  
GENERAL INFORMATION:  
APPLICANT: BAUM, Christopher  
APPLICANT: STOCKING-HARBERS, Carol  
APPLICANT: OSTERTAG, Wolfram  
TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF  
TITLE OF INVENTION: FOR GENE TRANSFER  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/793,610  
FILING DATE: 07-MAR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 44 31 973, 8  
FILING DATE: 08-SEP-1994

```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE 195 03 952.1
: FILING DATE: 07-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP95/03175
: FILING DATE: 10-AUG-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Berman, Richard J.
: REGISTRATION NUMBER: 39,105
: REFERENCE/DOCKET NUMBER: P1614-7007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)638-5000
: TELEFAX: (202)638-4810
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6505 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: DNA
: US-08-793-610-5

Query Match      25.4%; Score 512.8; DB 2; Length 6505;
Best Local Similarity 62.0%; Pred. No. 4,2e-143;
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;

QY 162 caggtgacccgtgattatgttggaatagtggttcgcttcgttatttggtaacataca 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2152 CAGGTAGCCCTATTATTATACAGTGGATTGCTGGGCTGCTGCTGCTTACATTCA 2211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 222 gatttccttggtgattatactgacgacgacagacccaagagagattcgaaaacagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2212 GGTTCATTATTGGGCGCTGGAGCGTGGAAACAATACAAAATTAACAAAGATTTTT 2271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 282 tcaatcagtttggcacagagacatcgctggttgatagctgtgacatcggtgaactaa 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2272 TCATGCTTAATATGCGACAGAGATAGCGTGGTGTGATGTGACAGCATGTTGGGAGACTTAA 2331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 342 caccgcgcgtgacacacattgac-----aaatacgaatgattgttggagaataagatgcct 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2332 CACCGGACTTACAGATGATGCTCTTAAGATTATGAACTTATGTTGGTACAAATTTGAAT 2391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 gtgtttcaaaacatgctactcttctcgatcgccctgacgttgggttggtgaagggctg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2392 GTTCTTTAGCAATAGCAACATTTTTCAGTGGTATTAGAGATTATACAGTGGTGG 2451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 459 gaaactaacctagtgactctataccaagtcctctctataatgcttcaaggcgacgactg 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2452 GAACCTAACCCCTTGATTTGGCCATCAGTCCTGTTGACATGTCAGCTGCTGTG 2511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 519 ttctagagtgatcatctatctgacgagtaagaatlaagtcctatactccaagctggggc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2512 GGCAGAAAGTACTACTTATTACTGATTAAGAACTCTTAAAGCAAAAGCTGGAGC 2571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 579 tgtggcagaagaagctctgtcatcaatccgaacagctcacttaaggccccagaagaa 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2572 AGTAGCTGAAGAGGCTTGGGAGCAATTAAGAACTGTGATTTGGAGGACAAAGAA 2631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 639 agaactcaaaagtccttccctttaataatacaagaatagctgtgttatttcccca 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2632 AGAAGCTTA----- 2640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 699 gtggactactaagtgtgtctgttnttgtaagfatacagaatctcaagaatgacaa 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2641 -----AAGGTAAACAAAAATTTAAGAAAGAGCTTA 2670
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 759 ggaattggcataaaagaactatagcttcaaaagtctctgtgtgtgtacttctt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2671 AAGATTGGGATTAAGAAAGCATATACAGCCAAATTTCTATAGTGTGCTTCTGTGT 2730
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 819 tatgaatgaaactatgactgtcttcttggtaaggaaacctctgattctaatggaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

DB 2731 GATCTATGATCTTAATGCTTGCGCTTCTGCTATGGAACACCTTGCTCTTAAGGGA 2790
QY 879 acctgatataccatcggaactgttctgtcttcttctttagtgaatccataagcagttia 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2791 -----AATTCATTTGGACAAAGTACTCAGTATTTCTTTCTGTATTAATTTGGGGCTTT 2844
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 939 ttgaattggagacgaatccctccacttgaacacttcgaatagcccgaggagctgtcct 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2845 TAGGTTGGACAGCATCTCCAGCATTTGAAGCAATTTGCAATGCAAGAGGAGAGCTTA 2904
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 999 tcatatttccagttattgttaagaaacccagatataacttttccacagctggata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2905 TGAATCTTCAAGATTAATTTGAATTAAGCCAAAGTATTACAGCTATTTCAGAGAGTGSCA 2964
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1059 taaacctgataccatagagaagacgtgtgaattttaaanaatgttcttcaattccalc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2965 CAACCCAGATTAATTAATTAAGGGAATTTGGAATTCGAATGTTCACTTCAGTTACCCATC 3024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1119 aagaccatctacaaagattcgaagagtcgaatccgaatlaagtlcgtgagagacgt 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3025 TCGAANAAGAGTTAAGATCTTGAAGGCGCTGAACCTGAAGGTGACAGAGTGGGACAGCT 3084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1179 cgccttggtgtgtcctcaatgacgtgaggaagatagcagtaagtcagcttctgcagaggt 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3085 GGCCTGTGTTGGAACCACTGCTGTGGGAAGACCAACAGCTCACTGATGACAGAGCT 3144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1239 atatgaccgagatgagtgcttatactatggttgatagaatgacatcaagactttaaactg 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3145 CTATACCCCAACAGAGGGGATGTGTCAGTGTGATGACAGAGATATTAGACATTAATGT 3204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1299 ggcgcattatcgagacatattgagtggttaagtaagaagcctgttctgtcggagcaac 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3205 AAGGTTCTACGGGGAATCATATTGCTGTGTGTCAGAGAACCTGATTGTTGTCACACAC 3264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1359 catcagtaacatatacaagtttgacagagatgattgacatgataagagatgagagaagc 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3265 GATAGCTAATAACATTCCTATGAGCGGTGAATAATGTCACATGATGATGAGAAAGC 3324
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1419 agcaagggaagcaaatgcatgatttcatcatagagattccataaatttaatacatt 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3325 TGTCAAGGAGAGCAATGCTATGACTTATCATGAACTGCCTCATTAATTGACACCT 3384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1479 gtlaggggaaaaaggagctcaaatgagtgagggcagagaacaaagagatcgcaattgtcgt 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3385 GGTGAGAGAGAGAGGGGCCAGTTGATGAGTGGCGCAGAGAGAGATGCGCATTCGACG 3444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1539 tgccttaattgaaaccccaagatctgattttagatagagctagctgtccctgagctc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3445 TGCCCTGTGTCGAACCCCAAGATCTCTGCTGATGAGGCGACAGCTGACGCTTGACAC 3504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1599 agaagaagatcagctgttcaagctgcacgtgagaag 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3505 AGAAGCGCAGACAGTGTTCACGTGCTGTGATGAAG 3542
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-08-793-610-6
: Sequence 6, Application US/08793610
: Patent No. 5858744
: GENERAL INFORMATION:
: APPLICANT: BAUM, Christopher
: APPLICANT: STOKING-HARBERS, Carol
: APPLICANT: OSTERAG, Wolfram
: TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEREOF
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Nikolaïdo, Marmelstein, Murray & Oram LLP
: STREET: 655 Fifteenth Street N.W. Suite 330
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005-5701
```







```
FILING DATE: 05-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/332,444
FILING DATE: 31-OCT-1994
APPLICATION NUMBER: 07/887,712
FILING DATE: 22-MAY-1992
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 4669 bases
TYPE: nucleic acid
STRANDEDNESS: singular
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: Genomic DNA
US-08-583-276-18

Query Match      25.1%; Score 506.4; DB 2; Length 4669;
Best Local Similarity 61.7%; Pred. No. 2.9e-141;
Matches 912; Conservative 0; Mismatches 476; Indels 90; Gaps 3;

OY 162 caggttgacccctgtatctgtgtgaatagtgctgtcctgtatcttgtgtatataca 221
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 760 CAGGTATGCCCTATTTATTCAGTGAATTCGTCGGGCTGCTGCTGCTTACATTCA 819
OY 222 gatttcccttggatataactgcagcagcagacgaagagattcgaacagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 820 GATTTCATTTTGGTCCCTGGCGAGCTGGAAACAAATACACAAATTTAGAAACAGTTT 879
OY 282 tcaatcagtttggcagcagcagcagcagcagcagcagcagcagcagcagcagc 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 880 TCATCTCTATATGACACAGAGATAGCTGTTTATGTCACGATGTTGGGACGTTAA 939
OY 342 cactgcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 940 CACCCGACCTTACAGATGATGCTCTTAAGATTATGACTTATGGTGACAAATTTGGAAT 999
OY 399 gttgttcaaaaatgtctacttcttgcattgagcctgagcagttgttctgtaaggctg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1000 GTTCTTACGTCATGCAATGAGCAACTTTTTCACCTGGTTTATAGTAGATTTTACACGCTGTTG 1059
OY 459 gaaactcccccagtgactctcatcagcgtctcctcttataatgcttcaaggcagcagc 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1060 GAGCTTAACCCCTTGATGATTTTGGCCATCAGTCTCTTCTTGAGCTGTCAGCTGCTG 1119
OY 519 ttctagaatgtcatcctcatgtacgaatgaagtaagtgcctatcccaagcctggagc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1120 GGCAAAAGATACATCTTCTATTTAGTAAGAACTTTAGCGGTATGCAAAAGCTGGAGC 1179
OY 579 tctggcagaagaagctctgtcatcaatccgaagcagcagcagcagcagcagcagc 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1180 AGTACTCTAAGAGGCTCTGGCGAGCAATTGAACCTGTGATTTGGAGGAGCAAAAGAA 1239
OY 639 agaacttcaaaagctcttcccttataataatacaagataatgtgtttatttccca 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1240 AGAAGCTTGA----- 1248
OY 699 gtgctaactaagtgctgtcttnttltgaagatacacagaatctcaaatgacaa 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1249 -----AAGTACAAACAAATTTTGAAGAAGCTAA 1278
OY 759 ggatttggcataaaagaactatagcttcaaaagtctcttctgtgtgtgttacttct 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1279 AAGAAATTTGGATTAAGAAAGCTATTACAGCCAAATTTCTATAGGTGCTGCTTCTGCT 1338
OY 819 tatgaatgaactatgactgtctgttcttgtatgaactcctcctgattcttaagggaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1339 GATCTATGATCTTATGCTCTGCGCTTCTGGTATGGAGCACCTTGGTCTCTCAAGGGGA 1398
OY 879 acctgatatacacacagcagcagcagcagcagcagcagcagcagcagcagcagc 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1399 -----ATATCTATTTGGACAAAGTACTCATCTATTTCTTGTGATTAATTTGGGGCTTT 1452
```

```
OY 939 ttgcattggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1453 TAGTTTGGACAGGACATCTCAAGCATTTGCAAGATTTTGAATTTGCAAGGAGGACCTTA 1512
OY 999 tcatatttccagctatgtatgaagaacccagatataactcttccacagcaggata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1513 TGAATCTTCAAGATTAATTGATTAATAGCCAGATTTGACACTTTTCAAGAGTGGA 1572
OY 1059 taaactgaatccatagaagaagcagcagcagcagcagcagcagcagcagcagc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1573 CAACACAGATTAATTAAGGAAATTTGAAATTTGAAATTTGCACTTACCTACCATC 1632
OY 1119 aagaccatctcaagatcttgaagagctgtgatactcgaatgaatgaatgctgggagaacgt 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1633 TCGAAAGAAGTTAAGATCTTGAAGGCTTGAACCTGGAAGGTCAGAGAGGGAACAGGT 1692
OY 1179 cgcccttgctcggtcccaatggcagtgaggagagcagcagcagcagcagcagc 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1693 CCCCCCTGTTGGAACAGTGGCTGTGGAGAGCACACACAGTCCAGCTATGCAGAGCT 1752
OY 1239 atatgaccgagatgagcagcagcagcagcagcagcagcagcagcagcagcagc 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1753 CTATGACCCACAGAGGAGGATGCTCAGTGTTCATGACAGATATTAGGACCATTAATGT 1812
OY 1299 gcgagcatatcgagacatattgagagtggttagtcaagagcctgttcttgggagcaac 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1813 AAGGTTTTCACGGGAATTCATTTGTTGTTGAGTCAGGAACCTGTAATTTTGGCCACAC 1872
OY 1359 catcagtaacaatcaatgaatgagcagcagcagcagcagcagcagcagcagc 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1873 GATACTGTAACCAATTCGCTATGCGCGTGAATAATGTCACATGATGATTAAGAAAC 1932
OY 1419 agcaaggaagcaaatgcagcagcagcagcagcagcagcagcagcagcagcagc 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1933 TGTCAAGAGAGCAATGCTATGACTTATCATGAACCTGCTCAATTTTGAACACCT 1992
OY 1479 ggttaagggaagaagagctcaatagagtgagagggcagaagaagagatctgcagctcgc 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 1993 GGTGGAGAGAGAGGCGCCAGTGTGAGTGGTGGCGAAGACAGAGATGCGCATTTGACAG 2052
OY 1539 tgccttagtgcgaaccccaagctctgatttgaatgagcagcagcagcagcagcagc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2053 TGCCCTGTTGCCAACCCCAATATCTCTCTGATGAGGCGCAGCTGACCTTGGACAC 2112
OY 1599 aagaagcaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 2113 AGAAAGCGAAGCAGTGTTCAGGTGAGCTGTGATTAAG 2150

RESULT 7
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```

APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36,677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-1

```

```

Query Match      24.9%   Score 503.4   DB 2:   Length 4264:
Best Local Similarity 61.2%   Pred. No. 2.2e-140;
Matches 902; Conservative 0; Mismatches 486; Indels 87; Gaps 2;

```

```

Db 162 caggctgaccctglatatgltggaataagtgctgctgcttgaatttggltacataca 221
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
477 CAGGTAGGCCATATATACATGATGATGCTGGGCTGCTGGTGTGCTTCAATTCA 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 222 gattcctctggtgattataactgacgacgacgacgacgacgacgacgacgacgacgac 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 537 GGTTCATTGTTGGCCCTGGCAGCTGGAACAAATACAAATTAACAAATTAACAAATTAAC 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 282 tcatcagcttggcacagacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 597 TCATGCTATATGCGACAGGAGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 342 cactcgacatgacacatgacacacacacacacacacacacacacacacacacacacacac 401
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 657 CACCGACTATACAGATGATGCTCCAGATTAATGAAGAAATGGTGAACAAATTTGAAAT 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 402 gtccaacaactgctactcttctgcatgctgctgacgacgacgacgacgacgacgacgacgac 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 717 GTTTCAGTCAATGCGCAATTTTTCACGCGTTATAGTAGAATTTACACGCTGGTTGAA 776
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 462 actacccatgagctatcaccgctcctccttataatgacgacgacgacgacgacgacgacgac 521
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 777 GCTAACCTTGTGATTTGGGCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 522 taggatgcatcctcatgaccagtaagaaatgaagtcctattccaaagctgagcctgt 581
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 837 AAAGTACTATCTTCACTTACTTAAAGAACTTAAAGCAATGCAAAAGCTGGAGCAGT 896
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 582 ggcagaagaatcttgatcataatccagacgacgacgacgacgacgacgacgacgacgacgac 641
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 897 AGCTGAAGAGCTCTGGCAGCAATTAAGCAATGATTCATTTGGAGGACAAAGAAAGGA 956
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 642 actccaagagcttccctttaataatacaagaatctggtttattttcccaagtg 701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 957 ACTTGA----- 962
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 702 gctactaagtgtgtctgtnttntgtaagglatcacagaatcctcaagaatgcaaga 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 963 -----AAGGTACAAACAAATTTAGAAAGCTAAAG 995
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 762 ttgtggcataaaggactatagcttcaaaagtctctctgtgtgtgtgtgtgtgtgtgtgtgt 821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 996 AATGGGATTAAGAAAGCTATTACAGCCAAATTTCTAAGTGTCTGCTTCTGCTGCTGAT 1055
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 822 gaatggaaactatgactgttltgtgtatggaacctctgattcttcaatggaac 881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1056 CTATGCACTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 882 tggatataccatcgagactgtctgtctgtcttctttagtgaatcatcagatgacgtatgt 941
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 1113 ---ATATTCATTGACACAGTACTGATTTCTTGTGATTAATATGCGCTTTAG 1169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 942 catggagcagcagctccctcacttggaaaccttgcgaatgcccaggaagctgcttca 1001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1170 TGTGGACAGCATCTCAAGCATTAAGCATTTGGCAAAATGCAAGACAGCAGCTTATGA 1229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1002 tatttccaggtatgataagaacccaagatagataacttcccaagctgataata 1061
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1230 AATCTTCAAGTATTTGATTAATTAAGCCAGATTTGACAGCATTTGCAAGGTGGCCAA 1289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1062 accctgaatcatalagaagaaactgtggaatttaaaatglttcttcaatatacaag 1121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1290 ACCAGATTAATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCACTTCACTTCCG 1349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1122 accatcatcaagatcttgaagaagctcgaatctcgaatgaatgaatgaatgaatgaatgaat 1181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1350 AAAAGAGTTAAGATCTTGAAGGCTGTAACCTGAAGTGCAGATGCGCAGACGAGTGGC 1409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1182 ctgtgcgtctcaatgagcagtggaagagtagcgtagtcagctctcgtgagaggtata 1241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1410 CCTGTTGGAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGTGAATGACAGGCTCTA 1469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1242 tgatcggatgagtgcttcatcgtgtgagatgagaaatgacatcagagccttaagtgcg 1301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1470 TGACCCACAGAGGAGATGTCAGTGTGATGAGCAGGATTTAGCAATTAATTTAG 1529
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1302 gcatatcagacatatttggatggttgaatgacagagcctgttltgtggagacacat 1361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1350 GTTTCAGGGGAATTCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1362 cagtaacaatalcaagatagacagatgagatgagatgagatgagatgagatgagatgagat 1421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1590 AGCTGAACAACTTGTGATGATGCTGGAAGATGTCACCATGATGATGATGATGATGATGAT 1649
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1422 aaggaagcaaatgcatgatttcatcagatgagatgagatgagatgagatgagatgagat 1481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1650 CAAGGAAGCAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1482 aggggaaagaaagagctcaaatgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1541
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1710 TGGAGAGAGAGGGGCCCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1769
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1542 cttagtgcgaaccccaagatctgattttagatgagtgacgtgctgctgctgacgacgacgac 1601
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1770 CCTGTTGCAACCCCAAGATCTCTGCTGATGAGGCCACGTCAGCCTTGACACAGA 1829
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1602 aagcaagcagctgttcaagctgacactgaggaag 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1830 AAGCGAACACTGTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1864
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

## RESULT 8

```

US-08-784-649A-5
Sequence 5, Application US/08784649A
Patent No. 5830697
GENERAL INFORMATION:
APPLICANT: Sikic, Branimir I
TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson
STREET: 2200 Sand Hill Road
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,649A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: Reg.No. 5830697 36, 677
REFERENCE/DOCKET NUMBER: 06037/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4264 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-784-649A-5

Query Match      24.8%; Score 501.8; DB 2; Length 4264;
Best Local Similarity 61.1%; Pred. No. 6.5e-140;
Matches 901; Conservative 0; Mismatches 487; Indels 87; Gaps 2;
```

```
882  tggatataccatcgagactgttcttgcgtttctttagtgaatccatagcaagtattcg 941
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1113  ---ATATTCCTATTGGACAGACAGTACTGATCTATTCTTTCTGTATTAATTTGGGCGCTTTAG 1169

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  942  catlgaagcagcagtcctccctcatttgaaccttcgcaatagcccgaggagctgccttca 1001
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1170  TGTGGACAGGCAATCTCCACAGATTAAGCATTTTGCATTAATGCAAGAGGACAGCTTATGA 1229

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1002  tatttccaggtattgataagaacccaagatagataacttcttccaaagctcgtgataa 1061
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1230  AATCTTCAAGATTAATTGATTAATTAACCAAGTATTACAGCTATTTCGAAGAGGGGACAA 1289

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1062  accitgaatccatagagaagcagctggaatttcaaaatttcttccattcaatccaaag 1121
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1290  ACCAGATTAATTAAGGAAATTTGAAATTCAGAAATGTTACTCTTACCTACCTACCTAC 1349

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1122  accatctataagatcttgaagagctgcaatcagaatgaatgaatcgtgagagacagtcgc 1181
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1350  AAAAGAGTTAAGATCTTGAAGGCGCTGAACTGAAAGTGAAGTGGGACAGCGTGGC 1409

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1182  ctgtgcggtcctcaatgagcaglyggaagagtaagcagtaagcagcttcgcaagagttata 1241
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1410  CCTGGTTGAAACAGTGGCTGTGGGAGAGACACACAGCTCAGTGAAGAGGCTCTA 1469

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1242  tgatccgagatgagcttattatgattgagatgagatgacatcagagctttaaattgctg 1301
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1470  TGACCCACAGAGGGGATGCTAGTGTGATGACAGAGATTAATGACCATTAATGTAAG 1529

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1302  gcaattacagagccatattgagtgtagtcaagagccctgttctgcggaaccacat 1361
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1530  GTTTCACGGGAAATCATTTGGTGTGTGATGATGAGAACCTGTATTTGCCACACAGAT 1589

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1362  cagtaacaatcagatlgagcagagatgagtgatcagatgaagaagatgagagagcagc 1421
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1590  AGCTGAAACATTCCTCTATGCGCCGTGAAATGTCCACCATGATGATGAGAAAGAGCTGT 1649

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1422  aaggaagcaaatgctgattatattatcatgagatgcttctaataaatttaacatttgt 1481
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1650  CAAGGAGCCAAATGCTTATGACTTATCAATGAATGCTCTTAATTAATGACACCTGCT 1709

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1482  aggggaaaaagagctcaaatagtgagagcagagaacagagatcgcacattgctgcg 1541
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1710  TGGAGAGAGAGGGGCCCACTTGAATGTCGCGAGAGACAGAGATGCCATTGCCACGTGC 1769

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1542  cttagctgaaaccccaagatctgatttagatgaggtacgctgcgtccctgattcaga 1601
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1770  CCTGGTTGCAACCCCAAGATCTCCTGCTGATGAGGCCACAGCTGACCTTGACACAGA 1829

      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy  1602  aagcaagtcagctgttcaagctgcacatgagagaag 1636
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db  1830  AAGCGAAGCAGTGTTCAGGTGCTGTGATTAAG 1864
```

```
RESULT 9
US-09-120-513-1
Sequence 1, Application US/09120513
Patent No. 6025160
GENERAL INFORMATION:
APPLICANT: Brun, Kimberly
APPLICANT: Chenery, Richard
APPLICANT: Ellens, Hanna
APPLICANT: Field, John
APPLICANT: Yue, Lin
TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
TITLE OF INVENTION: SEQUENCES ENCODING RAT MD1LB2 AND
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
STREET: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
```

```

; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GP50008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-09-120-513-1

Query Match      23.68; Score 477.4; DB 3; Length 4233;
Best Local Similarity 60.4%; Pred. No. 1.3e-132;
Matches 895; Conservative 0; Mismatches 496; Indels 90; Gaps 3;

162  caagtgacgacgtatgatgtggaataggtgtgctgccttgatttttgatcacata 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
355  catgtacgacctactatattacacgggcatgttgccggctgctctatgcttgcctacatcca 414
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
222  gatttccttgatgattataactgacgacgacgacccaagaagatttcgaaaaacagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415  gggttaccttgggtgctgcgacgtggagacaaatattacaaagattttagcagaagttttt 474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282  tcaatcagtttggacaagaacatcgctggttgatagctgtgacatcgtaactaa 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
475  ccattgccatcatgatgaatgagagatgagctggttgcgtgaatgacgtcgggagctcaa 534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
342  cactcgcattgaca---gacattgacaaatcatcagtgatgattggagataagattgctct 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
535  caccggcctcacagatagactctccaaatttaattgacggaatttgatgacaaacttggaaat 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
399  gtgtttcaaaaacatgctactcttttcgattgacctgacgttgagtttggtgaaagggctg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
595  gttctttcagtcctaaacgacatttttcagccgggttttatatagattttatagattggtg 654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
459  gaaactaacctagtagctataccacgctctcctctataatagcttcacggcggaacatg 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
655  gaaccttaacctttgatttttggccgcacagcccttatttgggtgctgcatctgcatgtg 714
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
519  ttccagatgtgatacctcattgacagtaagaattgaatgctcattccaaagctggggc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
715  ggcataaagctactgacttatttattataggaactccagccttattggaagacttgagc 774
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
579  tgtgacagaagaagctctgtcatcatccacgacagcatatagcctttagggccagaagaa 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
775  agttggcgaagaagacttttagcagccatcagaaactgtgatttcgtttggagacaaagaa 834
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
639  agaactcaaaagctcttctctttaaataaacaagatagtctgtgttlttlttccca 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
835  ggaacttga----- 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
699  gtggctactaagtgtgttctgttntttgtaaggtatcacagaatcotaagaatgcaaa 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
844  -----aaggtacatataaattttagaaagagcttaa 873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 759  gatattggcataaaaagaactatagcttcaaaaagtctctgtgtgctgtacttct 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 874  aagattggcatataaagaaagccatcacggccaaacatttccatagattgacctgtt 933
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 819  tatgaatgaacctgaagacttcttctgtataggaacctcttgatcttaatgaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 934  ggtcttattgctgtttatgacacttgcatcttgatgagacctcttgcgtct-----ctc 987
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 879  acctgatatcacatcgggactgttctgtcttcttctttagtgaatccatagaagta 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 988  aattgatatattttatggacaaagcttaccgcttcttcttcttcttatttttggggacttt 1047
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 939  ttgcatltggaagcagcagtcctcctactttgaaccttcgcaatagcccgaggagctt 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1048  cactatttgacatttgaccctcaaaacatagagccttggcaaatgacagggcgaccctta 1107
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 999  tcatatttccagtgattatgataagaaccagatagataacttccacagctgata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1108  tgaattcttcagatatttgaattatgataagccatgacattgacagcttctcaaccagagaca 1167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1059  taacctgaatccatagaagaagactgtggaatttaaaaatlytttcttcaatataccatc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1168  caaacacagacagatattatgggaatttggaaatttataaaatgtttacttcaactaacccatc 1227
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1119  aagaacatctatacaagatttctgaagtgctgaatctcagaattgaatcgtggaagacgt 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1228  accgaagtgaagttgaagcttcttgaaggccctcaacacgtgaagctgaagcggcgacgggt 1287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1179  cgccctgtgctgcgtcaatgtgacgttggaagagtagcgtatgcagcttcgtcagaagatt 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1288  agcccttggttggacacagtgctgtgaggaacacacacactgctcactgctcagagagcct 1347
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1239  atatgctcggatgatgttcttcatgtgtgataagaaatgacatcaagactttaaagt 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1348  ctacagaccctatagagggcgagctcagatcagcagcagacatcagacacatcagctcaatgt 1407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1299  gcggaattalcgagacacatatltggagtgttaagaaagccgttttgttggagccac 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1408  gaggatctcctgggaaatcatttgggtgtgagctcagaaacccgtgctgttccacacac 1467
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1359  catcagtaacatatacaatgaatgagacagatgaatgacatgacatgacatgacatgacatgac 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1468  gatttgcgaaacatttgcgtatgctgacgagaaacgtacatgagatgagatgagaaagc 1527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1419  agcaagaagaaacatgtcgatatttataatgagatttccataaattatacat 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1528  tgcataagaaacacatctgacttcatcattgaaactcccccacaaatttaaacacct 1587
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1479  ggtaggggaaaaagagctcaaatgagtgagggcagaagaagagtgcaattgctg 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1588  ggttggtgagagagagggcgacgtgagtgaggggacagaaacagagatcccatattgccc 1647
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1539  tgccttagtcgaaccacacagatctcgaattttagatgagctacgtcctcgtgattc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1648  ggccttgctgcgacaccccaagatccttggttgtagatgagggcagcagccttggagcac 1707
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1599  agaaagcaatcagctgttcaagctgagagagata 1639
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1708  agaaagcgaagccgtggttcagccgcttgaataaaggcta 1748
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10.
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chienery
; APPLICANT: Harma Eilens
; APPLICANT: John Anthony Feld
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF

```

FILE REFERENCE: GP-50008-D1  
CURRENT APPLICATION NUMBER: US/09/450,105  
CURRENT FILING DATE: 1999-11-29  
EARLIER APPLICATION NUMBER: 09/120,513  
EARLIER FILING DATE: 1998-07-22  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 4233  
TYPE: DNA  
ORGANISM: HOMO SAPIENS  
US-09-450-105-1

Query Match 23.6%; Score 477.4; DB 4; Length 4233;  
Best Local Similarity 60.4%; Pred. No. 1.3e-132;  
Matches 895; Conservative 0; Mismatches 496; Indels 90; Gaps 3;

QY 162 caggttgaccctgtatattggtggaatagtgctgctgattttgtgtatata 221  
DB 355 catgtacgctactatatacagggcatgtgctggtgtgcttccatcgcttcaaccca 414  
QY 222 gatttcctgttgatataacatcgcagacagacagacagagatcgcgaacagtttt 281  
DB 415 ggtttcacttgtgctgagcgtggtgagacataacacagatttagcagaagtctt 474  
QY 282 tcatcagtttggcagacagacatcgctggtttgatatgcttgatcgtggaactaa 341  
DB 475 catgacacatcagatcagagatagctgtgttgatgacgtgagggagctcaca 534  
QY 342 caatcgacatgaca---gacatgacacaaatcagtgatgtatgagataagatgctc 398  
DB 535 ccccgccgacacagatgagctgtcccaaaatatacggaattgtgacaaaacttggat 594  
QY 399 gtgtttcaaaacatgtctactcttttgcattgctgctgagctgtgttgtaaggctg 458  
DB 595 gtcttctcagctataacagacatttccagcgttttataatagattataagtgttg 654  
QY 459 gaaactcaacctagtgacctctatccagctctcctcttaataatggtctcagcgagcatg 518  
DB 655 gaagctaaacctgttaatttggccgtcagccctctattgtgtgtgtcactgtgtg 714  
QY 519 tcttagagtgatcatcatctatgacacagtaaggaatagtgcttcttccaaactgagc 578  
DB 715 ggcacaaagtactgactctcttcaataaagaacccaggtcttaagcgaagctggagc 774  
QY 579 tgtggcagaagaagctgtgtgtcaatccgaacagctcatagctctttagggccagagaa 638  
DB 775 agttgcggaagaagctgttagcagccatcagaactgtgattgcttggaggaacaaaga 834  
QY 639 agaacttcaaaagctcttctctttaaataacaagaatagctgtgtttatttcccca 698  
DB 835 ggaacttga----- 843  
QY 699 gtggtactaagtgtgtctgttnttltgaaagtatacagaagaatcctaaagtgcmaa 758  
DB 844 -----aaggtacataaataattgaagaagctaa 873  
QY 759 ggaatttggcataaaagaagataagcttcaaaagtctgtgtgtgtgtactctt 818  
DB 874 aagaatttggcataaagaagacatcagcgaacattctcatagatttagtgcctactgt 933  
QY 819 tatgaaatgaacctatgactgtcttttggataggaacctcctgattcctaaatgaga 878  
DB 934 ggtctatgcttctatgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 987  
QY 879 acctgatatcatcaggaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 938  
DB 988 aatgataattctatgtgacagtgcttctctctctctctctctctctctctctctct 1047  
QY 939 ttgcatgtgagcagcagcttccctcactttgaaaccttcgaatagcccgagagctgct 998  
DB 1048 cagtatgtgacatttagccccaacatagaaagccttctgcaaatgcaagggcgagccta 1107

QY 999 tcatatttccaggtatattgataaagaaccagtatagataacttttccacagcttgata 1058  
DB 1108 tgaatcttccaagataattgatacagcgaagcagcttcttccaaagagata 1167  
QY 1059 taacctgatatcagaagaagctgtggaatttaaaatgttcttcaattatcacc 1118  
DB 1168 caaacagacagatataatggaatatttgaattttaaataatgttacttcaatcacc 1227  
QY 1119 aagacatctacaaagatctgaaagtgtcgaatcagaatlaagctgagaagacag 1178  
DB 1228 acgaagtgaagttaagatcttgaaggtcctcaacctgaagtggaagagcgagcagcgt 1287  
QY 1179 cgcctgtgtcgttctcacaatgagcagtggaagtgatcaggtccagctctcagaagct 1238  
DB 1288 agccctgtgtgcaaacagtgctgttggaagaagcaaacgtgtcgtctgcaagagct 1347  
QY 1239 atatgaccgatatgattgcttcaatcgtgtgtgataagtgacatcagaagcttlaagt 1298  
DB 1348 ctacgaaccatagagggcgaggtcagatctgcagacagacatcagaacatcagat 1407  
QY 1299 gcgcatatcagacacatattggaatggttagtcaagaagcctgtttgttcggagcac 1358  
DB 1408 gagtatctgcgggaatcattcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1467  
QY 1359 catcagtaacatataatgaatgagcagatgattgactgtatgaagaatggaagagc 1418  
DB 1468 gatgcccgaatacatcgtatcgttcgagcaaaacgtcacacatgagatgagaagaagc 1527  
QY 1419 agcaaggaagaacatcgtatcatttcaatgagcttccataatataatcatt 1478  
DB 1528 tgtcaaggaagaacatcgtatcatttcaatgagcttccataatataatcatt 1587  
QY 1479 ggttaagggaagaagcgtcaaatgagtggaagggcagaacagagatcgcgaattgctg 1538  
DB 1588 ggtgtgtgagagagggcgagctgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1647  
QY 1539 tgcctagtctgaacacccaagatctgtattttagatgtggtcagctgcgacctgagctc 1598  
DB 1648 ggcctgtgtcgcgaacccaagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1707  
QY 1599 agaaagcaagcagctgttcaagctgacgtgaggaagata 1639  
DB 1708 agaaagcgaagcgtgtgttcaagcgtcgtgtgataagcta 1748

RESULT 11  
US-08-461-823-1  
Sequence 1, Application US/08461823  
Patent No. 553840  
GENERAL INFORMATION:  
APPLICANT: Bhattacharya, Satish K.  
APPLICANT: George Jr., Albert L.  
APPLICANT: Nazarenko, Irina  
TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OncorPharm, Inc.  
STREET: 200 Perry Parkway  
CITY: Gathersburg  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20877  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,823  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 08/168,621
FILING DATE: 16-DEC-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/010,433
FILING DATE: 27-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kaita, Glenn E.
REGISTRATION NUMBER: 30,649
REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301 527-2058
TELEFAX: 301 208-6997
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-461-823-1

```

```

Query Match      11.0%; Score 221.8; DB 1; Length 2726;
Best Local Similarity 59.3%; Pred. No. 2.7e-56;
Matches 398; Conservative 0; Mismatches 267; Indels 6; Gaps 1;

```

```

972 ctccgaatagcccgagagctgccttccatattcccaagttatgataagaaccag 1031
1495 CTATGCCAAGCCAAATATCAGACGCCACATCATCATCATCATCATCATCATCAT 1554
1032 tatagataactttcccaagctgataataaaccgaatccatagaaagaaactggaatt 1091
1555 GATTGACAGCTACAGCGACGAGGCTTAATGCCAACAACATTGGAAGAAATGTCACATT 1614
1092 taataatgtttcttcaatataatcatcaagaacatctatcaagaattctgaaggctgaa 1151
1615 TGGTGAAGTTGTATTCAATCAATCCACCGACCGACATCCCATGCTTCAGGAGCATGAG 1674
1152 tctcaagaatgaagctcgagagagagagctgccttgcgtctcaatgagcagtggaagag 1211
1675 CTTGGAGGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1734
1212 tacgtagtccagcttctgcgagagagctatataatgacccgagatgagcttcatcagtgatg 1271
1735 CACAGTGTGCTCAGCTCTGAGAGCGGTTCTACGACCCCTGGCAGAGGAAAGTGTCTTGA 1794
1272 tgaagaatgacatcagagcttaaatgtgcgagcattatcgagacatattggaagtgttag 1331
1795 TGGCAAGAAATAAAGGAGACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1854
1332 tcaagaagcgtgtttgttcgagaccacatcagtaacaatatcaagtatgaa-----cg 1385
1855 CCAGAGACCCCATCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1914
1386 agatgagtgatcgtatgaagaagatgagagagagagagagagagagagagagagagagag 1445
1915 CCGGCTGTGTGTCACAGAGAGAGATGCTGAGGCGAGCAAGAGAGGCGCAACATACATGCTTC 1974
1446 tatcatgagagtttccataaatttaacatatacattgtgtaggggaaaaagagagctcaatgag 1505
1975 CATGAGTCACTGCTTATATAATATAGACATTAAGTAGAGAACAAAGAACTCAGCTCTC 2034
1506 tggagggagagaacagagatcgcaatgtcgtgcttagttcgaaaaccccaagattct 1565
2035 TGGTGGCCAGAAACAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094
1566 gattttagatagctagctgctgccttgagattcagaagaagagagagagagagagagagagag 1625
2095 GCTTTTGGATGAAGCCACGCTGAGCTTGATGATGAGAAAGTGAAGAAAGTTTGTCCAGAGAGC 2154
1626 actggaagaag 1636

```

DB 2155 CCGGAGCAAG 2165

RESULT 12

US-08-996-545-1  
Sequence 1, Application US/08996545

Patent No. 592898

GENERAL INFORMATION:

APPLICANT: Skatrud, Paul L.

APPLICANT: de Maard, Maarten A.

APPLICANT: Peery, Robert B.

TITLE OF INVENTION: Multiple Drug Resistance Gene atnd of

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/996,545

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-11766

TELEPHONE: 317-276-3334

TELEFAX: 317-276-2763

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4002 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..4002

US-08-996-545-1

Query Match

Best Local Similarity 8.7%; Score 175.2; DB 2; Length 4002;

Matches 421; Conservative 0; Mismatches 383; Indels 6; Gaps 1;

```

839 ttgttttggtagaacctccttgaattcttaatgagaagacgtgatataccatggga 898
3020 TTGGCTGCGGTTTGGTAGCGAGGAGACTCTTGTGTCACCAAGATGACATTTTCC 3079
899 ctgtcttgcgtttctttagtgtaatccatagcagttatgtgagcagcagctcc 958
3080 GCTTCTTTGTTGTTTCCGAGATTCCTTTGGTGTCAATCGGGGCGACGCTTTT 3139
959 ctcaattgaaaccttgcgaatagccgagagagctgccttcatatttccaggtttg 1018
3140 CTTTGCACGACATGGGCAAGCGAAGAAATGCGCGCGCGCAATTCGAGAGCTGTTCG 3199
1019 ataagaaccagatagatatttccacagcttgatataaaccgaatcagatagaag 1078
3200 ACCGAAAGCCCAAAATGATGCTGCTGTAAGAGGCGGAGAACTCGAAAGGCTGGAAG 3259

```

QY 1079 gaactgtggaatttaaaatgtttcttcaattatccatcaagaccatctatcaaatc 1138  
DB 3260 GGGAAATTCGAATTTAGGACGTGCACCTCAGATACCGACCCGCGCAAGACCTGTCC 3319  
QY 1139 tgaagtcgtgaatcagaatcaagtcgtgagagacagtcgcctgtgctgtcctcaag 1198  
DB 3320 TCGCGCGCTTGACCTGACCGCGTGAACCCCTGACAAATATTTGCGCTTGTGCGACCCAGG 3379  
QY 1199 gaagtggagagagtagtgatccagctctcgaagaggtatataatgccgagtgatgct 1258  
DB 3380 GTTGTGCGAAGAGTACACCATTTGCTGAGCGCTTTTACGATTCGATTCGATTCGCGGCT 3439  
QY 1259 tctatgtgtgagagagtagcatcagagctttaagtgtgagcttctcgaaccata 1318  
DB 3440 CCAATCTTTGTAAGGAGACATAGTAACATAATATCAATCTTACCGACCTTTC 3499  
QY 1319 ttgagagtgtagtcaagagcctgtttgtcgtgagaccacatcaagtaacaataatcaag 1378  
DB 3500 TGTCACTGTGTCAGCGACGACGACACTTACAGGGGACCATCAAGAAACATTTTAC 3559  
QY 1379 atgagagc-----agatgtgtgactgtatgagagatgagagagcagaggaagcaaa 1432  
DB 3560 TGTGATTTGTCGAAGATGACGACCGAAGAAATTTGATTAAGGCTTGCAAGACGCTA 3619  
QY 1433 atgcgatgtatttatcatgagagttccctaataattatcatgtgaggggaaaag 1492  
DB 3620 ATATCTACAGATTCATCATGCTGCTCCGAGGGCTTTATATACAGTTGTCGACGACAAAG 3679  
QY 1493 gagctcaaatgagtgagagcgagaaacagagagatcgcaatgtcgtgcttgaattcgaa 1552  
DB 3680 GAGGCGATGTTGCTGCGCGCCCAAAACGACGTGTGSCCATTTGCCGCCCTTTCTTCGGG 3739  
QY 1553 accccaagattctgatttagatgaggtcagctcgtgcctgattcagaagaagcaagtcag 1612  
DB 3740 ATCCCAAAATCCCTTCTTCGATGAGAGCGACGACCTCGACCTCGAGTCAAGAAAG 3799  
QY 1613 ctgttcaagctgacgtgaggaagataccc 1642  
DB 3800 TCGTCCAGCGCGCTTGTGATGCCGCTGCC 3829

RESULT 13  
US-08-966-545-3  
Sequence 3, Application US/08996545  
Patent No. 5928898  
GENERAL INFORMATION:  
APPLICANT: Skatrud, Paul L.  
APPLICANT: de Maard, Maarten A.  
APPLICANT: Peery, Robert B.  
TITLE OF INVENTION: Multiple Drug Resistance Gene atyrd of  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: U.S.  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,545  
FILING DATE:  
CLASSIFICATION: A35  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Thomas D.  
REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-11766  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3334  
TELEFAX: 317-276-3334  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4002 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-966-545-3

Query Match 8.7%; Score 175.2; DB 2; Length 4002;  
Best Local Similarity 37.7%; Pred. No. 2.9e-42;  
Matches 305; Conservative 116; Mismatches 383; Indels 6; Gaps 1;

QY 839 ttgcttttggatggaacccctcctgattccttaatgagagaaactgatataccatcgga 898  
DB 3020 UUGCGCTCGGCUUUGGUAACGAGGAGCACUUCUUGGUCACCGAGUAUGACAUUUGCC 3079  
QY 899 ctgtcttctgttttctttagtgaatccatagcagatgattgagagagcagctcc 958  
DB 3080 GCUUCUUGUUGUUCUUCGAGAUUCUUCUUGGUCUUAUCCGGGCGACCGUUCUUU 3139  
QY 959 ctcaacttgaacactcgcgaatgaatgaatgagagagctgcttcaattatctcagatgatt 1018  
DB 3140 CCUUGSCACAGACAGUUGGCGAAGGCGAAGAUUGCGCGCCGCGCAUUCGACAGUUGG 3199  
QY 1019 ataagaaccacagatagataaacttccacagcttgatataaactgataatcagataag 1078  
DB 3200 ACCGAAGCCACAAUUGUUAACUGGUCUGAAGAGGCGGAGGUCGAAAGGUGGAG 3259  
QY 1079 gaactgtggaatttaaaatgtttcttcaattatccatcaagaccatctatcaaatc 1138  
DB 3260 GGGAAATTCGAATTTAGGACGTGCACCTCAGATACCGACCCGCGCAAGACCTGTCC 3319  
QY 1139 tgaagtcgtgaatcagaatcaagtcgtgagagacagtcgcctgtgctgtcctcaag 1198  
DB 3320 TCGCGCGCTTGACCTGACCGCGTGAACCCCTGACAAATATTTGCGCTTGTGCGACCCAGG 3379  
QY 1199 gaagtggagagagtagtgatccagctctcgaagaggtatataatgccgagtgatgct 1258  
DB 3380 GTTGTGCGAAGAGTACACCATTTGCTGAGCGCTTTTACGATTCGATTCGCGGCT 3439  
QY 1259 tctatgtgtgagagagtagcatcagagctttaagtgtgagcttctcgaaccata 1318  
DB 3440 CCAATCTTTGTAAGGAGACATAGTAACATAATATCAATCTTACCGACCTTTC 3499  
QY 1433 atgcgatgtatttatcatgagagttccctaataattatcatgtgaggggaaaag 1492  
DB 3620 ATATCTACAGATTCATCATGCTGCTCCGAGGGCTTTATATACAGTTGTCGACGACAAAG 3679  
QY 1493 gagctcaaatgagtgagagcgagaaacagagagatcgcaatgtcgtgcttgaattcgaa 1552  
DB 3680 GAGGCGATGTTGCTGCGCGCCCAAAACGACGTGTGSCCATTTGCCGCCCTTTCTTCGGG 3739  
QY 1553 accccaagattctgatttagatgaggtcagctcgtgcctgattcagaagaagcaagtcag 1612  
DB 3740 ATCCCAAAATCCCTTCTTCGATGAGAGCGACGACCTCGACCTCGAGTCAAGAAAG 3799  
QY 1613 ctgttcaagctgacgtgaggaagataccc 1642









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:32:28 ; Search time 9392.76 seconds  
(without alignments)  
2312.123 Million cell updates/sec

Title: US-09-873-409-15  
Perfect score: 2021  
Sequence: 1 ttccgccttgctgctgact.....taaaatgctgactttatgga 2021

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estipl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estrov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	545	27.0	545	10	AL040762
2	489.8	24.2	533	11	N24315
3	489.8	24.2	533	11	N24315
4	482.8	23.9	559	11	BF692596
5	385.8	19.1	405	10	AA243820
6	378.8	18.7	872	12	AK020318
7	296.2	14.7	894	12	BF584668
8	233.6	11.6	780	11	BF587938
9	232	11.5	503	13	AK0380214
10	221.8	11.0	943	10	AL520322
11	219.8	10.9	661	11	BF302675
12	219.4	10.9	932	13	AZ670821

Result No.	Score	Query Match	Length	DB ID	Description
13	218.6	10.8	880	13	AZ687805
14	213.8	10.6	575	10	AM342644
15	209.2	10.4	886	13	AZ540627
16	209	10.3	853	13	AZ679807
17	206.4	10.2	891	11	BF298756
18	200.8	9.9	803	13	AZ682250
19	199.8	9.9	787	11	BF584063
20	197.8	9.8	402	11	BF82550
21	196	9.7	886	13	BF139685
22	191.2	9.5	871	13	AZ682350
23	189	9.4	947	13	AZ683753
24	188	9.3	1019	11	BF248052
25	187.8	9.3	885	13	AZ686798
26	187.6	9.3	860	13	AZ543293
27	185	9.2	750	11	BF585786
28	184.8	9.1	611	11	BF635924
29	184.6	9.1	687	11	BF646725
30	184.4	9.1	834	13	AZ548312
31	183.6	9.1	897	13	AZ541090
32	181.2	9.0	656	10	AM686102
33	178.8	8.8	823	13	AZ532602
34	177.8	8.8	878	13	AZ539524
35	176.2	8.7	651	10	BE016688
36	172.4	8.5	921	13	AZ687628
37	167.6	8.3	795	11	BF313560
38	167.4	8.3	919	13	AZ690701
39	166.8	8.3	944	11	BF796582
40	166.2	8.2	785	10	AV709991
41	164.4	8.1	815	11	BF409057
42	164.4	8.1	2676	12	AK014319
43	163.4	8.1	842	13	AZ671925
44	161.2	8.0	541	10	AZ722858
45	160.6	7.9	899	13	AZ548799

#### ALIGNMENTS

RESULT 1  
AL040762 545 bp mRNA EST 29-FEB-2000  
LOCUS DKFP2434C1815\_r1 434 (synonym: htes3) Homo sapiens cDNA clone  
DEFINITION DKFP2434C1815 5', mRNA sequence.  
ACCESSION AL040762  
VERSION AL040762.1 GI:5409708  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 545)  
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.  
EST (Bloecker, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Bloecker H

#### FEATURES

Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de, Braunschweig/Germany) within the cDNA sequencing consortium of the Human Genome Project.  
s1 sequence also available.  
This clone (DKFP2434C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY. Email: clone@rzpd.de.  
Location/Qualifiers  
1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFP2434C1815"



```

OY 1975 tctaagaaatcgaggtctttttttaaagctgacttatgga 2021
DB 481 CCTAAGAAAATCGACAGCTTC-TTTTAAATGCTGCTTATGGA 526

RESULT 3
LOCUS U66692 533 bp mRNA HTC 23-JUL-2001
DEFINITION Homo sapiens clone EST422562 mRNA sequence.
ACCESSION U66692
VERSION U66692.1 GI:1906577
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 533)
TITLE Allikmets, R., Gerrard, B., Hutchinson, A. and Dean, M.
JOURNAL Characterization of the human ABC superfamily: isolation and
MEDLINE mapping of 21 new genes using the expressed sequence tags database
REFERENCE Hum. Mol. Genet. 5 (10), 1649-1655 (1996)
97049974
2 (bases 1 to 533)
AUTHORS Allikmets, R., Gerrard, B. and Dean, M.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1996) Human Genetics Section, National Cancer
Institute, NCI-FCRDC, Frederick, MD 21702, USA
FEATURES
source 1..533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="EST422562"
/note="similar to ATP-binding cassette transporter"
BASE COUNT 139 a 134 c 118 g 142 t
ORIGIN

Query Match 24.2%; Score 489.8; DB 12; Length 533;
Best Local Similarity 98.1%; Pred. No. 2,7e-120;
Matches 517; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

OY 1497 tcaaatgagtgagggcagaagaagatcgcaattgctgctgactgactgcaaaccc 1556
DB 1 TCAAAATGATGGAGGGCAGAAACAGAGATCGCAATTGCTCGGCTTACTTGAAGAACCC 60

OY 1557 caaagattcgatttagatgaggtacgtctgcccctgattcagaagaagaagtcagctgt 1616
DB 61 CAAGATTCTGATTTTATGATGAGGCTACGTCTGCCCTGGATTTCAGAAAGCAAGTCAGCTGT 120

OY 1617 tcaagctgacacgaggaagatacccccaagctattcatttgaacctaatccactcaag 1676
DB 121 TCAAGCTGACACGAGAAAGATACCCCGAGGATTTTTCACCTAATTTTCACCTCAAG 180

OY 1677 tggagaatcgctgaccttgaaccagcgccctcgacagctctggccctcaaacctcacc 1736
DB 181 TGGAGAATCGCTGACCTTAAACCAAGCGCCCTTCGACAGCTCTGGCCCTCAACCTCACC 240

OY 1737 ctgacctcctgctgacctgactgactgacataaccatcaaggccatgcaagtgtgagc 1796
DB 241 CTGACCTCCTGCTGCTGCTATGAGCTACTGACATACCTCAAGGCCAATGCAAGTTGGGCC 300

OY 1797 ctgacccaataatacactgatacttagagggaggttgagcaagtggcgatagtaaacacatt 1856
DB 301 CTGACCAAAATTAACATGTAATAGAGGGGAGTTGGCGGTGCGGTATGAAAAAACCAATT 360

OY 1857 gaagaattcttcgagtgagctgactccctataaaccagagccttaagagcccttaca- 1915
DB 361 GAAGATTTTCTCGATGGCTGACTCCCTTATTAACCAAGCCTTACAGCCCTTACAAA 420

OY 1916 -ggcttaatgacacatttacttgcatttgccttgaagtgagtaagcgttttttttc 1974
DB 421 GCCTTAATGAGGACATTTACTTTGCAATTGCTTGAAGATGATGAAGGTTTATTTTTC 480

```

```

OY 1975 tctaagaaatcgaggtctttttttaaagctgacttatgga 2021
DB 481 CCTAAGAAAATCGACAGCTTC-TTTTAAATGCTGCTTATGGA 526

RESULT 4
LOCUS BF692596 559 bp mRNA EST 22-DEC-2000
DEFINITION 602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
ACCESSION BF692596
VERSION BF692596
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE 1 (bases 1 to 559)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DP
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1203 row: a column: 13
High quality sequence stop: 555.
FEATURES
source 1..559
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4334100"
/clone_lib="NIH_MGC_62"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (p1 phage-resistant)"
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgagctcgagc); Site_2: SfiI (ggcgatgagc);
Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows:
5' adaptor sequence: 5'-ATCTGAGAGCGCGAGCGCGGACGACATG-dT(30)BN-3'
(where B = A, C, G or N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

```

BASE COUNT 161 a 118 c 141 g 139 t
ORIGIN

Query Match 23.9%; Score 482.8; DB 11; Length 559;
Best Local Similarity 96.8%; Pred. No. 2e-118;
Matches 546; Conservative 0; Mismatches 12; Indels 6; Gaps 5;

OY 1240 tatgacccgagatgagttatcatcagtgagtgagatgacatcgagagtttaagt 1299
DB 1 TATGATCCCGATGATGCGTTTATACCGTGGATGAGATACATCAGACCTTTAAATGTG 60

OY 1300 cggcattatcgaaacatatltgagtgagtgagtgagagcgtgttgttcggagacc 1359
DB 61 CGGCATTATGCAACACATA-TGGAGTGGTTAGTCAAGAGCTG-TTTGTTGGGAGCACACC 118

OY 1360 atcagtaacaataatcaagtagagagagatgagtagatgagtagagagagagga 1419
DB 119 ATCAGTAAACAAATATCAAGATGACGACGAGATGATGATCAATCAAGATGAGAGCA 178

OY 1420 gcaagggaagcaatcgtatgattatcatcagtgagtttccataaatatatacatg 1479

```

Db 179 GCAAGGAGCAATGCATATGATTTATCATGACGTTTCTTAATTAATTAATCATATG 238  
 QY 1480 gtaggggaaaa-aggagctcaaatgagtgagggcagaagaagagatcgcaattgctcg 1538  
 Db 239 GTAGGGGAAAAACAGAGCTCAATGAGTGAGGCGCAGAAACAGAGGATCCGAATTTGCTCG 298  
 QY 1539 tgccttagtgcgaaccccaagatctgatttagatgaggtcgcgtcgcctggatc 1598  
 Db 299 TGCCTTAGTTGGAACCCCAAGATTCTGATTTAGATGAGCTAGTCTGCCCTGGATTTC 358  
 QY 1599 aagaagaagctcagcttctcaagctcagctcagagagatccccaggtatcatttg 1658  
 Db 359 AGAAGGAGAGCAGCTGTCAGCTGACCTGAGAGAGATACCCCAAGATTCATTTTG 418  
 QY 1659 accaatctcaccctcaagtgagatcgctgaccttgaccagcgcccttgacagctct 1718  
 Db 419 ACCAATTTTACCTCAGTGAATGAGATGCTGACCTTGACACAGGCCCTTGACAGCTCT 478  
 QY 1719 ggcctcacaacctaaccctgacctctgctgctctatgactgacatcacatcaagg 1778  
 Db 479 GGCCCTCAAC--TCACTGACCTCTGCTGACTGATGAGTACATACCTCAAG 536  
 QY 1779 ccatacgagttgagccctgac 1802  
 Db 537 -CATATGAGTTGTGTGGCTGAC 559  
 RESULT 5  
 AA243820 405 bp mRNA EST 06-AUG-1997  
 LOCUS z167906.1 Soares\_NHMPu.S1 Homo sapiens cDNA clone IMAGE:668506 5'  
 DEFINITION similar to SW:MDRI\_MOUSE\_P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;  
 mRNA sequence.  
 AA243820  
 VERSION AA243820.1 GI:1874631  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 405)  
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,  
 Scheinberg,K., Steptoe,M., Tan,F., Theisling,B., White,Y., Wyllie,  
 T., Waterston,R. and Wilson,R.  
 Mashu-Merck EST Project 1997  
 COMMENT Unpublished (1997)  
 CONTACT Wilson RK  
 WASHINGTON UNIVERSITY School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 THIS clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1572 Std Error: 0.00  
 Seq primer: -28ml3 rev2 ET from Amersham.  
 FEATURES  
 source  
 1. 405  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5562473"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:668506"  
 /clone\_lib="Soares\_NHMPu.S1"  
 /tissue-type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10b"  
 /note="Organ: mixed (see below): Vector: p7773D-Pac  
 (pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NbM, pregnant uterus  
 NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."

BASE COUNT 115 a 81 c 85 g 124 t  
 ORIGIN

Query Match 19.1%; Score 385.8; DB 10; Length 405;  
 Best Local Similarity 99.5%; Pred. No. 1.8e-92;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 903 tcttggtttctttagtgaatcacaagagatgcatcttgagagcagtcctca 962  
 Db 17 TCTCTGTTCTTTCTTTAGTGAATCATACAGATTATGATGAGCAGCAGTCCCTCA 76  
 QY 963 cttggaaccttcgcaatagcccgagagcgcttccttatatttcaggtattgataa 1022  
 Db 77 CTTGGAACCTTCGCAATAGCCCGAGAGCTGCCCTTCAATATTTTCAGATTATGATTA 136  
 QY 1023 gaaccacagatagataactttccacagctcgatataaactgatacacaagaagac 1082  
 Db 137 GAACCCAGTATAGATTAATCTTCCACAGCTGAGATATACTGATCATAGAGAAC 196  
 QY 1083 tctggaattaaaaatgtttcttcaattatcacaagacatctcaaatctgaa 1142  
 Db 197 TGTGGAATTTAAAAATGTTCTTCAATTAATCATACAGACCAATCATCAAGATTCTGAA 256  
 QY 1143 aggtctgaatctcgaatgaatgctcggagagaacgctgctgtgctcgaatggag 1202  
 Db 257 AGGTCTGAATCTCAGAAATTAAGTGTGAGAGACAGTCCCTTGGTGTCTCAATGGCAG 316  
 QY 1203 tgggaagatcagcgtgagccagctctcagagagtgatataatccagatgagcttat 1262  
 Db 317 TGGGAAGATCAGTACGATGACAGCTTCTGACAGGTATATATATCGGATGATGGTTTAT 376  
 QY 1263 catggtgagatgagaatgacatcagagctt 1291  
 Db 377 CATGCTGATGATGAGATGACATCAGAGCTT 405

RESULT 6  
 AK020318 872 bp mRNA HTC 05-JUL-2001  
 LOCUS AK020318  
 DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched  
 library, clone:9230106F14, full insert sequence.  
 ACCESSION AK020318  
 VERSION AK020318.1 GI:12860872  
 KEYWORDS CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,  
 clone.lib:RIKEN full-length enriched mouse cDNA library  
 clone:9230106F14.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 872)  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Methods in enzymology. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE 2 (bases 1 to 872)  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome research. 10 (10), 1617-1630 (2000)  
 MEDLINE 11042159  
 PUBMED 11042159  
 REFERENCE 3 (bases 1 to 872)  
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,D., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,







```

OY 1064 ctgaatccatagaaggaactgtaattcaaaatgttcttcaattatccaaagac 1123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 TAGAGACACTTACTGACTGTTGAACTGAAAATATGCACTTCTTATCCATCAAGAC 183
OY 1124 calctatcaagatcttgaagaagcttgaatctcaaatgaatgaagagacagtcgct 1183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 CTAAGATCTGTGAATCCCAATGATTTCTCTTGAGTGTCTCTCGCGAAGACATAGCTT 243
OY 1184 tggctggtctcaatgcaagtggaagaagtaacggtgaagctcttcgagaagttatag 1243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 TGGTGTGTGACAGCGCTCTGCGAAGACATGCTGTTCTCTTATTTAGAGATCTATG 303
OY 1244 atccggaatgatgcttcatcatcagtgatgagaatgacatcagagctttaaatgtagc 1303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 ATCCAACTTCAGACACAAATGATTTGGATGGCATGACATTAATCACTTGAACTTAAT 363
OY 1304 attatcgaagcattatgagtgatgtagtgaagagcctgttctgtcggagaccacata 1363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 GGTGAGGCAACAAATAGAGACTAGTGAGCCAAAGACCTGTTGTTGCCACACGATTC 423
OY 1364 gtaacaatctcaagtgagagagatgagtgactgataagaagatgagagagagcaag 1423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 GAGAAATATTAATCTCTTGGAGAGCCGATGCAAAACAGGTGAGATGGAAGAGCTGCTA 483
OY 1424 ggaagcaaatgctgtagtatttatacagagagttcttaataatttaacattgtag 1483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 GGGTTCCTAATGCTCTATTCATTCATCATCAAACTTCTGAGGCTTTGAACCTCAGGTG 543
OY 1484 ggaagaaagagctcaaatgagtgagagagagagagagagatcgcagtcgctgtc 1543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 GAGAAAGAGAGACTACAACTTTCTGAGAGCAAAACAAAGATAGCAATGCAAGGCA 603
OY 1544 tagtctgaagaccccaagatctgattttagatgagagctacgctcgtcgtgattcaga 1603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 TGTATAAATATCAAGCAATTTCTTCTCTGATGAGGCAACAGTGCATTTGACTGTAAT 663
OY 1604 gcaagtcagctgttcaagctgcacitga 1631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 CAGAAAAGCTGTGTGCAAGAGCACTTGA 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
LOCUS A0380214 503 bp DNA 29-JUN-1999
DEFINITION RPCI-11-161L23.TV RPCI-11 Homo sapiens genomic clone RPCI-11-161L23
, DNA sequence.
ACCESSION A0380214
VERSION A0380214.1 GI:4351237
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 503)
Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter
,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Other GSSs: RPCI-11-161L23.TJ
Contact: Shao, Zhao, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@edj.med.bufileo.edu). Clones may be purchased from
BACpac Resources (http://bacpac.med.bufileo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

```

```

FEATURES
    source
        Seq primer: T7
        Class: BAC ends.
        Location/Qualifiers
            1..503
                /organism="Homo sapiens"
                /db_xref="GDB:7561726"
                /db_xref="taxon:9606"
                /clone="RPCI-11-161L23"
                /clone_1ib="RPCI-11"
                /sex="Male"
                /cell_type="Lymphocytes"
                /note="Vector: PBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPCI11 Human Male BAC library"
BASE COUNT 183 a 93 c 86 g 139 t 2 others
ORIGIN
Query Match 11.5%; Score 232; DB 13; Length 503;
Best Local Similarity 98.8%; Pred. No. 3.1e-51;
Matches 243; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
OY 666 tataaagaatgcttggtgttatttccacagtgactaagtggtgtgtntt 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 TATAACNAGATATGCTGTTTATTTTCNCAGTGGCTACTAAGTGTGTTCTGTT-TT 445
OY 726 tgaagatatacagaatctcaagaatgcaagagatttgagataaaggagctatagc 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 444 TGTAAAGTATACACAGATCTCAAGATGCAAGATGCAAGATTTGGATMAAAGACTATAGC 385
OY 786 ttcaaaagtgctcttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 TTCAAAAGTCTCTTGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 325
OY 846 ttggtatggaacctcttgaattcttaatgagaacctgatalaacaatcggaactgtct 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 324 TTGCTATGGAACCTCCTGATCTTATGATGAGAAGACCTGATATACATCGGACTGTCT 265
OY 906 tgctgt 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 TGCTGT 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
LOCUS AL520322 943 bp mRNA 13-FEB-2001
DEFINITION AL520322 LVL_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
prime, mRNA sequence.
ACCESSION AL520322
VERSION AL520322.1 GI:12783815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 943)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr; Web : www.genoscope.cns.fr.
Location/Qualifiers
    1..943
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CS0DB006YC15"
        /clone_1ib="LVL_NFL004_NBC2"
        /sex="male"
        /tissue_type="neuroblastoma cells"
        /lab_host="DH10B"
        /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
        was primed with a NotI-oligo(dT) primer. Five prime end

```

EST.  
zebrafish.  
Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Rasbora; Danio.  
1 (bases 1 to 651) x Johnson S. Li F. Marra M. Eddy S.

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: S.L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

QY 1032 tatagataactttccacagctgatatnaacctgataccatagaagaacgttggaatt 1091

Sequencing by: Washington University Genome Sequencing Center Clone  
distribution information can be found through the I.M.A.G.E.

150 GATGATACACACAGGAAAGCCATATATCCAGGACGATATGGAGAGATATCTGAT 102  
1092 TAAAGTGTCTTCAATATCATCAAGACATCTATCAAGATTCTGAAGTCTGAA 1151

Trace considered overall poor  
Seq primer: T7 from Gibco  
High quality sequence stop: 1

DB 190 TGGTGAAGCTGATATTCACATATCCACCCGACCGGACATCCCGATCCTTCAGGAGCAGTGG 243

QY 1152 tctcagaattagctctggagagacagtcgcctgtctcgctctcaatgycagctgggaagag 1211

inos

```
1. .661
/organism="Danio rerio"
/strain="AB"
```

[illegible]

```
/clone="3815766"  
/clone_lib="Sugano Kawakami zebrafish DRA"  
/sex="mixed (one male and one female, including
```

1272 *tgagaa*tgacatcagagccttaatgtgcggaattatcgagacatatggagtggttag 1331

```

/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/note="Vector: PME18S-FL3; Site_1: DraIII (CACTGCTGTG);

```

QY 1332 tcaagagcgttttctgtagccacatcaagtaacaatatcaagtgtga-----cg 1385

with an oligo(dT) primer [ATGTCGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTTCGGCCTACTG], digested and cloned into distinct DraIII

1386 agatgatgtcactgatgaagagatgtgagagacagcaaggaagcaaatgcgtatgattt 1445  
 | ||| | | ||||| | | | ||||| | | | ||| |  
 1456 agatgatgtcactgatgaagagatgtgagagacagcaaggaagcaaatgcgtatgattt 1515  
 | ||| | | ||||| | | | ||||| | | | ||| |

CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (Institute of Medical Science and University of Tokyo).

Oy 1446 tatcatgaggttccataaatltaatacatgtgtgggaagaagtccaatgag 1505  
||| | ||||||||| | | ||||| | |||||| |

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTCTCTCTTAAAAAGCTCG and 3' end primer CGACCTCAGCTGAGACA. "

1506 tggagggcagaacacagagatcgccaattgctgctttagtgcgaacccaagatttc 1561

ORIGIN

DB 11: Length 661.

Oy 1566 gatttagatagagctaactgtctgcctgattcagaagaacgaatcagctgttcaagctgc 1622

Best L  
Matche:

```

1 Similarity 59.0%; Pred. No. 6.le-48;
377; Conservative 0; Mismatches 262; Indels 0; Gaps 0

```

Qy 1626 actgagaag 1636  
| | | | | | |

Db 63

||||| || | ||||| |||||  
TGTGTTACTATTCCACGCAGATCGGCACCTTCCTTGGGTAGCAAGCAGCTCCAACACTGT 580  
  
||||| ||||| ++++++ggtatgaagaac 1027

RESULT 11

Db 57

AGGCCATTGACTAGGCCCGTGGTGCCTTATGAGGCTACAACTACTTGCATGCCCTC 520

LOCUS	661 bp	mRNA	EST	23-FEB-2001
DEFINITION	BG302675			
	fl35b04.x1	Sugano	Kawakami	zebrafish DRA Danio rerio cDNA clone
	3815766 3'	similar to TR:Q91586	Q91586	MULTIDRUG RESISTANCE

Db 51

[illegible]

ACCESSION	13100131
VERSION	BG302675.1
GI	13100131
Accession	BG302675
Version	BG302675.1
GI	13100131

Db 45

ATTCAAAACATTTATTTTCATTTACCCCTCCAGAAAGATGTGACGATTTTTCGAAGGAA 400



Y	1384	-----cgaatgatgctcgcattgaagaagatgagagagcagaaggaagcaat	1434
Db	429	GCTAGAGATGGAGAAACCAACCACTGAAAGACAGATGATGATGCTAAATGGCAAAAT	370
Y	1435	gcgatgatgttcatcagagagttccctcataaatttaacatggtgtaggggaaaagga	1494
Db	369	GCACATGATGATTTATTTTCATCTTCACAAAGAGATATGATTAATGTAGAGGAAAAAGGA	310
Y	1495	gctcaaatgagtgtagggcagaagaagagagatgcgaatgctgctgctgctgctgaac	1554
Db	309	GCTGCATTATCATCAGAGGACAAACAAAGAAATTCCTATTGACAGTGTGATTATGAAAA	250
Y	1555	cccaagatctgcattttagatgagctacgctcgcctcgtatcagaagaatcagct	1614
Db	249	CCTAACATTTTATTTACTTGATGAAAGCTACATCAGACCTTGATTACCAAAAGTAAAGATT	190
Y	1615	gttcaagctgcactgtaggaagata	1639
Db	189	GTACCAACAGACCTTGAAAAAGCTA	165
RESULT	14		
LOCUS	AM342644/c		
DEFINITION	AM342644 575 bp mRNA EST 31-JAN-2000		
ACCESSION	U010106.x1 Sugano Kawakami zebrafish DRB Danio rerio cDNA clone		
VERSION	264067.3, similar to TR:Q91586 Q91586 MULTIDRUG RESISTANCE		
KEYWORDS	PROTEIN, mRNA sequence.		
SOURCE	AM342644		
ORGANISM	AM342644.1 GI:6839010		
REFERENCE	EST.		
AUTHORS	zebrafish.		
COMMENT	Danio rerio		
TITLE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
JOURNAL	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;		
COMMENT	Cypriniformes; Cyprinidae; Rasbora; Danio.		
	1 (bases 1 to 575)		
	Sugano, S., Kawakami, K., Johnson, S., Li, F., Marra, M., Eddy, S.,		
	Hillier, L., Clifton, S., Allen, M., Gibbons, M., Jost, S., Kucaba, T.,		
	Martin, J., Pepe, D., Steptoe, M., Underwood, K., Theisling, B., Ritter		
	, E., Bowers, Y., Wylie, T., Waterston, R. and Wilson, R.		
	Washu zebrafish EST Project 1999		
	Unpublished (1999)		
	Contact: S.L. Johnson		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watsn.wustl.edu		
	Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA		
	Sequencing by: Washington University Genome Sequencing Center		
	Seq primer: T7 ET from Amersham		
	High quality sequence stop: 465.		
FEATURES	Location/Qualifiers		
Source	1..575		
	/organism="Danio rerio"		
	/strain="AB"		
	/db_xref="taxon:7955"		
	/clone="264067"		
	/clone.lib="Sugano Kawakami zebrafish DRB"		
	/sex="mixed (one male and one female, including		
	unfertilized eggs)"		
	/dev_stage="adult"		
	/lab_host="DH10B (phage resistant)"		
	/note="Vector: pME18-FL3; site.1: DraIII (CACTGTG);		
	site.2: DraIII (CACATG); 1st strand cDNA was primed		
	with an oligo(dT) primer (ATGCGCCCTTTTCTTTTCTTTT);		
	double-stranded cDNA was ligated to a DraIII adaptor		
	(TTTGGCTACTGG), digested and cloned into distinct DraIII		
	sites of the pME18-FL3 vector (5' site CACTGTG, 3' site		
	CACATG). XhoI should be used to isolate the cDNA		
	insert. Size selection was performed to exclude fragments		
	<1.5kb. Library constructed by Dr. Sumio Sugano		
	(University of Medical Science) and		





Db 562 ATTGAAACCTACATTTTATTACTTGATGAGCTACATCAGCACTTGATACACAAAGT 621  
Qy 1606 aagtcagctgttcaagctgcactgagagaagata 1639  
Db 622 GAAAAGATTGTACACACAGCACTTGAAAAAGCTA 655

Search completed: April 22, 2002, 18:32:35  
Job time: 12248 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 23:09:55 ; Search time 886.72 Seconds  
(without alignments)  
1954.004 Million cell updates/sec

Title: US-09-873-409-15

Perfect score: 2021  
Sequence: 1 ttccgcttgcgtgagact.....taaatgcgtacttatgga 2021

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1661242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_1101:\*

1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	533.6	26.4	4279	22	AAD03506	Dog P-glycoprotein
2	532	26.3	4279	22	AAD03504	Dog P-glycoprotein
3	532	26.3	4279	22	AAD03505	Dog P-glycoprotein
4	530.4	26.2	4279	22	AAD03488	Dog P-glycoprotein
5	528.8	26.2	4317	22	AAD03489	Dog P-glycoprotein
6	512.8	25.4	3860	22	AAZ49332	Human wild-type mu
7	512.8	25.4	4349	22	AAH57442	Human intestine ce
8	512.8	25.4	4646	15	AAQ72872	Human multdrug re
9	512.8	25.4	4646	21	AAZ94738	Human ATP binding
10	512.8	25.4	4669	19	AAV32645	Human P glycoprote
11	512.8	25.4	6505	17	AAT13394	Hybrid vector pSF-

12	512.8	25.4	8630	21	AAZ24042	Retroviral vector
13	511.2	25.3	3860	21	AAZ49333	Human G185V mutant
14	511.2	25.3	4669	8	AAH70752	Sequence of human
15	511.2	25.3	4669	14	AAQ52726	Sequence of human
16	511.2	25.3	8630	21	AAZ24041	Retroviral M4 mdr-
17	509.6	25.2	4378	11	AAO04522	Multidrug Resistan
18	504.8	25.0	4186	22	AAH6127	Cynomolgous monke
19	504.8	25.0	4195	22	AAH6128	Cynomolgous monke
20	503.4	24.9	4264	19	AAH65533	Mutated human P-gl
21	501.8	24.8	4264	19	AAH65534	Mutated human P-gl
22	498.4	24.7	3988	21	AAZ88973	Human MDR-1 DNA.
23	480.6	23.8	4189	21	AAZ49334	Murine multidrug r
24	480.6	23.8	4313	14	AAO38950	Mouse multidrug re
25	477.4	23.6	4233	21	AAZ90198	Rat mdr1b2 multir
26	477.4	23.6	4233	22	AAH27498	Rat mdr1b2 multir
27	477.4	23.6	4788	21	AAZ49335	Murine multidrug r
28	469.2	23.2	4425	21	AAZ52048	Rat multidrug resi
29	467.6	23.1	4369	21	AAZ52047	Rat multidrug resi
30	445.4	22.0	3924	21	AAZ94742	Human ATP binding
31	445.4	22.0	3924	21	AAZ88974	Human MDR-3 DNA.
32	349	17.3	4776	21	AAZ94744	Human ATP binding
33	222.6	11.0	4102	21	AAO45942	Arabidopsis thalia
34	221.8	11.0	2726	15	AAQ70907	Multidrug-resistan
35	221.8	11.0	2726	15	AAO70916	Multidrug-resistan
36	221.8	11.0	2726	18	AAT43322	Multidrug resistan
37	208	10.3	1810	20	AAV69393	H. contortus PGP-A
38	208	10.3	4175	20	AAV69392	H. contortus PGP-A
39	205.4	10.2	465	22	AAH38132	Probe #6818 used t
40	204	10.1	216	22	AAH51285	Probe #19971 used
41	199.6	9.9	3512	20	AAH69395	H. contortus PGP-O
42	192.4	9.5	14000	21	AAO90920	A. thaliana ATPAC
43	190.8	9.4	4051	21	AAO90919	A. thaliana ATPAC
44	182.6	9.0	1957	21	AAC35415	Arabidopsis thalia
45	177.8	8.8	6300	21	AAA96070	AtPgp-1 ATP bindin

#### ALIGNMENTS

RESULT 1	
AAD03506	
ID	AAD03506 standard; CDNA: 4279 BP.
XX	
AC	AAD03506;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Dog P-glycoprotein (PGP) allelic variant (Genotype D) CDNA.
XX	
KW	Dog: P-glycoprotein allelic variant; PGP: multidrug transporter;
KM	MDR1: drug bioavailability; transgenic animal; genetic model; ss.
XX	
OS	Canis familiaris.
XX	
PH	Key
FT	Location/Qualifiers
FT	17.3862
FT	CDS
FT	
FT	allele
FT	/*tag= a
FT	/product= "Dog P-glycoprotein (PGP) allelic variant
FT	(Genotype D) protein"
FT	replace (91, T)
FT	/*tag= b
FT	replace (607, C)
FT	allele
FT	/*tag= c
FT	replace (1001, T)
FT	/*tag= c
FT	replace (3458, A)
FT	/*tag= c
XX	
PN	WO200123540-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000MO-US26767.

PR	28-SEP-1999;	99US-0156510.
XX		
XX	(GENT-) GENTEST CORP.	
PA		
XX		
PI	Stocker PJ, Steinmel-cresp1 DT, Cresp1 CL, Relf TC, Patten CJ;	
XX		
DR	WPI: 2001-235373/24.	
XX	P-PSDB; AAE00310.	
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful	
PT	for determining the bioavailability of drugs and for screening for dog	
PT	PGP inhibitors -	
XX		
XX		
PS	Claim 9; Page 102-107; 11pp; English.	
XX		
CC	The invention relates to dog P-glycoprotein (PGP) also referred	
CC	as multidrug transporter (MDR1) and nucleic acids encoding them.	
CC	The invention also includes fragments and biologically functional	
CC	variants of dog P-glycoprotein. PGP and their nucleic acids are	
CC	useful for determining the bioavailability of drugs and for	
CC	screening PGP inhibitors. They are useful for the diagnosis and	
CC	treatment of conditions characterised by PGP activity, by	
CC	reducing or increasing PGP activity in a cell. PGP nucleic acids	
CC	are used as oligonucleotide probes, complements of PGP nucleic	
CC	acids are useful as antisense oligonucleotides, to induce a PGP	
CC	'knockout' phenotype. They are used to prepare a non-human	
CC	transgenic animal, which are valuable as genetic models for	
CC	human diseases.	
CC	The present sequence is dog P-glycoprotein (PGP) allelic variant	
CC	(Genotype D) cDNA. The PGP enzyme functions as an efflux pump	
CC	exporting small molecules across the cell membrane. This enzyme	
CC	is a member of the ABC transporter family.	
XX		
Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;		

Query Match	26.48;	Score 533.6;	DB 22;	Length 4279;
Best Local Similarity	62.98;	Pred. No. 1.7e-148;		
Matches 929; Conservative	0;	Mismatches 459;	Indels 90;	Gaps 3;

Oy	162	cagtttgaacccctgattatgtttggaaatagtggtgtgcgccttgatttttgataacaca	221
Db	358	cacgaaatgacattatcaacagtggaacgagtcggcgcgctgtgtgtgccttaacacca	417
Oy	222	gatttccttctgtgaataaactgcagcagcagacccaagaagatcttcgaacaacgtttt	281
Db	418	ggttcattctctgtgcctgtgcagcaggaagacagataactcaaatatgaanaacaatttc	477
Oy	282	tcatttcagttttgacacagacacalcgctgtgtttgatagcctgtgacattggtgaactaa	341
Db	478	tcatgtctatcatgcgacacagagatattgtctgtttgaacgttgatcaacgttggagactaa	537
Oy	342	cactgcacatgcacagacattg--acaaatcatagatcgttatctgaagatagaattgtct	398
Db	538	caccgggtctcacagacagatgctctccaatatcaatgaaggaatttgcgacaaatggaat	597
Oy	359	gttgtttaaacaacatgctactcttttcgatattgcccctggcagttggttgtgaaaggctg	458
Db	598	gtctcttcaatcaatagaaacatttttcaaccgggtttatagtgtaggtgttacaagtgtg	657
Oy	459	gaacttcaccttagtactatcatccacagctctccctctataatgtgcttcacgcgagcatg	518
Db	658	gaagctaaaccttgatattttggccalcacagcccgcttcctggaacttcaagccgcacatcg	717
Oy	519	ttcttagatgtgtcatctatttcagcagtaagaataatagttcctattccaacagcttggcg	578
Db	718	ggcaaaagtactatcttcatctattacttgataaagaactcttgcctatgcataaagcttgagc	777
Oy	579	tgtgtgcagaagaagctctgtcatcaatctcgaacagctacagctcttagggccacagaa	638
Db	778	agtagctgaagaagaagctcttagcagaacatacgaacacgttgatctcttgttggaggaacaaaga	837

QY	639	agaactctcaaaaggtccttccttctaataataaagatagtctgtttattttccca	698
Db	838	agaacttga-----	846
QY	699	gtgcgtactaagtgtgtctgttnttctgaagctatcacagaactctccaagtcaaa	758
Db	847	-----	876
QY	759	ggatttggcataaaaagactatagctctccaagaatgtctctgtgtgtatcctt	818
Db	877	aggaatttggataaagaagactatccacgccaacattctatttggcgcgtcttcatt	936
QY	819	tatgaatggaaacctatgactgtcttttttggatctggaacctctctgatcttaatggga	878
Db	937	gattcatatgcataatgtctctgtgtctgtgatactggacctctctgtgtctctccagta	996
QY	879	accttgatatccactcggaacttctctgtgtgtttcttcttagtgaatccataagcaatga	938
Db	997	-----atactctcttggacaagtaactcaatgtctctcttcttctgtatatttgggacctt	1050
QY	939	tgtcatttggagcaagcttccctcaacttctggaaccttgcgaataacgacccgagagctgtccct	998
Db	1051	tagtatttggacagcctaccaccaagcatttgaagcatttggacaacgcgaagagagcagctta	1110
QY	999	tcaattttccaggtctatttataagaaccacagataagataactttccacagctcgtgata	1058
Db	1111	tgaactcttccaagaataatttaccataaaacagacatttgcagactatttcgaagaagtgcaca	1170
QY	1059	taaaccttgataccatagaaagaaactcttggaatttaaaaatgtctcttccaattatccatc	1118
Db	1171	taaacagataataatttaaggaagaatttggatccaanaaatgttccacttcagttacccttc	1230
QY	1119	aagacacatctatacagaattcttgaagaaggtctcgaactctcagaataatgaactcttgagagacagat	1178
Db	1231	tcgaagaagaagttaaagttcttaagaaggtctccaacctgaaggtctcagagtctggcagagaagct	1290
QY	1179	cgctctgtgtcgtctcaatgtgcagctgtgggaagagtaacgtaacggtcttcgcagaggtt	1238
Db	1291	ggcgtctgttctgggaacagttgtctgtcgggaagagacacgcgtgcagctgtatgtcagaaggtct	1350
QY	1239	ataatgctccggaatgaatgtgctttatcagtggtgagctggaatctgacacagagctttaaatgt	1298
Db	1351	ctataagaccacagaatgtgcattgtctgtatgtatgtatgtgacagagccattgagaccataaatgt	1410
QY	1299	gcgcgatatcatcagaacatatatttggatgtgttgaatcagaagccctgttcttgcggagccac	1358
Db	1411	aaggaatcttcgggaataattactgtgtgtgtgagtgagtgcaaggccctgtgttgcacacac	1470
QY	1359	catcagtaacaatatcaagtatgtgaagcagatgtgtgtgactgaatgaagagatgtgagaagc	1418
Db	1471	gataccttgaaaaacatttgcataatgcgcggaanaatgtcacacatgagatgtgaatgaagaagc	1530
QY	1419	agcaagggaaagcaaatgcgatgatatttatcagatggggtcttccataataatttaataact	1478
Db	1531	tgtaagaagaagccaatgcactgtatatttatcagaagaactcaactataataatttgaacacct	1500
QY	1479	gttaagggaaaaaagagctcaaatgtgagtgagggcagaaacagagagatctgcgaatgtcgtg	1538
Db	1591	ggttggagagagaggggcgcacgctgaggtgtgttgacaagaacaagagatctgcgcatgtcgtg	1650
QY	1539	tgaccttagtctcgaaaaccaccaagatctgattatagatgagctgaactgtgcctcgtgattc	1598
Db	1651	ggccctgttctgcacaaccaccaagattcttctcgtcgtgacatgaggaacgtcagctcttgacac	1710
QY	1599	agaaagcaagtcagctgtttaaagctctgacatggaggaag	1656
Db	1711	tgaagatgaagcagctgttctcaagtggtgcctcgtgataag	1748

RESULT	2
AAD03504	
ID	AAD03504 standard; cDNA; 4279 BP.
XX	

AC AAD03504;  
 XX 13-JUN-2001 (first entry)  
 DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.  
 XX  
 XX Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;  
 KM drug bioavailability; transgenic animal; genetic model; ss.  
 XX  
 OS Canis familiaris.  
 XX  
 XX Key Location/Qualifiers  
 FT 17..3862  
 FT CDS /\*tag= a  
 FT /\*product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype A) protein"  
 FT replace (607, C)  
 FT allele /\*tag= b  
 FT  
 FT  
 XX MO200123540-A2.  
 XX  
 XX 05-APR-2001.  
 XX  
 XX 28-SEP-2000; 2000WO-US26767.  
 XX  
 XX 28-SEP-1999; 99US-0156510.  
 XX  
 XX (GENT-) GENEST CORP.  
 XX  
 XX Stocker PJ, Steimel-crespi DR, Crespi CL, Reif TC, Patten CJ;  
 XX  
 XX WPI. 2001-235373/24.  
 DR P-PSDB; AAE00308.  
 XX  
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX  
 XX Claim 9; Page 85-90; 11pp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 CC  
 XX Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Query Match 26.3%; Score 532; DB 22; Length 4279;  
 Best Local Similarity 62.8%; Pred. No. 5e-148;  
 Matches 928; Conservative 0; Mismatches 460; Indels 90; Gaps 3;

0Y 162 caggttgcccgctatgttggaatggttgctgctgatttgggtacataca 221  
 DB 358 cagcatgctatattacagtgagatcggtgctgctgctgctgctacatcca 417  
 0Y 222 gatttcctgtgataatactcagcagcagacagcaagagattcgaaacagtttt 281  
 DB 418 ggttcattctgtgctcgagcagcaggaagacagataccacaaattagaacaattttt 477

0Y 282 tcaattcagtttggcacagacatcgctggtttgatagcgtgacatcggtgaactaa 341  
 DB 478 tcaigtatcatgacagcagagatgctgtgttgaagtcgtacgttggtggagctaa 537  
 0Y 342 cactcgcatacagacacatgtt---acaaatcagltgagtgatgagataagattgctct 398  
 DB 538 caccgcgtcacagacagctgtctccaaatcatgaagaaattgtagcgaacaaattggaat 597  
 0Y 399 gttgtttcaaaacatgctactttttcgtatggtcctgagctgtgttggtaagagctg 458  
 DB 598 gtcttcatcaatagaacaattttccacggtttatagtggtggtttacacgtgttg 657  
 0Y 459 gaactcacccctagctactatccacgtctcctcttaatagttcgaagcgcagctg 518  
 DB 658 gaagctaaccttgatgttggccatcaagccctgttcttggaacttcaagccgcatctg 717  
 0Y 519 ttctagatggtcatcctatgacacagtaagaaatgagcttccaaagctgggagc 578  
 DB 718 ggcagaagatcatcttcatcttactgataaagaactcttgacctatgcaaaagctggagc 777  
 0Y 579 tgtgcagaagaagctctgtcatcaatccgaacagtcataagcctttagggccagagaa 638  
 DB 778 agtagctgaagaagctcttagcagcaatcagaactgtgattgcttggagagacaaagaa 837  
 0Y 639 agaacttcaaggtcttccctttaaataaacaagatatgttggtttatcttcccca 698  
 DB 838 agaacttga----- 846  
 0Y 699 gtgctactaagtgtgtctgttnttgaagtatcacagaatcctaagaatgcaaa 758  
 DB 847 -----aagtaacaaanaattagaagaagctaa 876  
 0Y 759 ggaatttgcataaagaagcataagctcaaaagctctctgtgctgtgtactctt 818  
 DB 877 aggaattgagataaagaagcctacacgccaacttctatgtgtgcgccttcttact 936  
 0Y 819 tatgaattgaacttgaactgctcttggatggaactcctgattcctaagagga 878  
 DB 937 gatcatcatcatatgctctgcttccgtgagatgagacccctcgtctccatcaga 996  
 0Y 879 acctgatatcaacacgagctcttctgtcttcttcttgaatccatagacagta 938  
 DB 997 -----atatcttatgacagaagctactcactgtctcttcttgaatcaatggggctt 1050  
 0Y 939 ttgcattgagcagcagctccctacttgaaaccttcgaacatgacccgagagagctgctt 998  
 DB 1051 tagtatgacagcatcccaagcatltaagcatttgaacgaagaggggagcagctta 1110  
 0Y 999 tcaattttccaggttatgataaagaacccagatagataactttccacagctggata 1058  
 DB 1111 tgaactctcaagaataattgacataaacaagcatltagcagctatlcgaagatggaca 1170  
 0Y 1059 taacctgaatcatagaaagcagctggaatttaaatgttcttccaatccatc 1118  
 DB 1171 taaccagataatataaggaagaatttgaattccaanaatgttcaactcgaattacacctt 1230  
 0Y 1119 aagaccatcatcaagatctgaaagcttgaatcagaatlaagctgtgagagacagt 1178  
 DB 1231 tcgaagaagaagtaagcttaagaggtctcaacctgaaggttcgaagtgagcagacagt 1290  
 0Y 1179 cgcttggtcggttccatagcagtgaggaagatcaggtagtcacagcttccagaagtt 1238  
 DB 1291 ggcgtgtgttggaacaagctgctgcgggaagcagcagccgctgagctgtagcagaagct 1350  
 0Y 1239 atatgacgcagatgagcttcatcatcagtgatagaaatgagacgaagctttaaagt 1298  
 DB 1351 ctatgaaccaacagatgagctgtctgtatgattagcagcgacatagaccataaagt 1410  
 0Y 1299 gcgcattatcagagacacatattggaatgagtgatcaagaagcctgttcttgaggacac 1358  
 DB 1411 aagcatctctcggaataattacgtgtgtgagtcagagcctgtgttcttgccacac 1470  
 0Y 1359 catcagtaacaatacaagtatcagcagagatgatgtgactgattgaagagatgtgagagac 1418

Db 1471 gtagctgtaaacatctgcatgcccgcgaataatgcacatgagatgagataagaagc 1530  
 QY 1419 agcaaggaggaacaaatgcatgatttcatgaggttctcctaataatttaacatt 1478  
 Db 1531 tcttaaggaggaacaaatgcatgatttcatgaggttctcctaataatttaacatt 1590  
 QY 1479 ggttaaggaggaacaaatgcatgatttcatgaggttctcctaataatttaacatt 1538  
 Db 1591 ggttggagagagagggccacagctggtgtgacagaacagagatcgccatgctcg 1650  
 QY 1539 tgccttagtgcgaaccccaagattctgatttagatgaggtcagctcgccctgattc 1598  
 Db 1651 ggcctcggttcgcaaccccaagattctctgctgagtaggcaacgctccttgacac 1710  
 QY 1599 agaaagcaagtcagctgctcaacgtcactgaggaagc 1636  
 Db 1711 tgaagtgaagcagtggttcaagtgccctgataagc 1748

## RESULT 3

AAD03505  
 ID AAD03505 standard; cDNA; 4279 BP.

AC AAD03505;

XX 13-JUN-2001 (first entry)

DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.

KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
 KM MDR1; drug bioavailability; transgenic animal; genetic model; ss.

XX Canis familiaris.

XX Location/Qualifiers

FT CDS 17..3862

FT /\*tag= a /product= "Dog P-glycoprotein (PGP) allelic variant  
 (Genotype B) protein"

FT allele replace (91, T)

FT /\*tag= b /tag= d  
 FT allele replace (607, C)

PN MO200123540-A2.

XX 05-APR-2001.

PD 28-SEP-2000; 2000MO-US26767.

PF 28-SEP-1999; 99US-0156510.

PR (GENT-) GENTEST CORP.

PI Stocker PJ, Steinmetz-crespi DT, Crespi CL, Reif TC, Patten CJ;

XX WPI: 2001-235373/24.

DR P-PSDB; AAE00309.

XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -

PS Claim 9; Page 93-99; 11pp; English.

XX The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by

CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.

XX Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Query Match 26.3%; Score 532; DB 22; Length 4279;

Best local Similarity 62.8%; Pred. No. 5e-148; Matches 928; Conservative 0; Mismatches 460; Indels 90; Gaps 3;

QY 162 caagttgacccctattatggtgataagtggtgctgctgattttggtacataca 221  
 Db 358 cactgtagcctattatggtgataagtggtgctgctgattttggtacataca 417  
 QY 222 gatttcctggtgattataaactgcacacagacagacaggaagattcgaataatttt 281  
 Db 418 gtttcaattctggtgctgacagagacagacagatactcaataattagaacaatttt 477  
 QY 282 tcatcagtttgacacagacacatcggtggttgatagctgtgacatcggtgaattaa 341  
 Db 478 tcatcagtttgacacagacacatcggtggttgatagctgtgacatcggtgaattaa 537  
 QY 342 cactgcagtcagacagacttg---acaaatcagtgatgattgagataagttgctc 398  
 Db 538 caccgcgctcacagacagatgctccaaatcagtgatgattgagataagttgctc 597  
 QY 399 gttgttcaaaactgctacttcttgcattgctgctgacgtggttggtagaaggctg 458  
 Db 598 gttcttcaatcaatagcaaatcttccacggttttataagtggttaccagctggtg 657  
 QY 459 gaaactaccctagtgactctatccacagctctcctctataatagctcagcgacatg 518  
 Db 658 gaagtaacccttggtgatttggccatcagcctgcttctggaattcagcgcacatcg 717  
 QY 519 ttctagatggtcattccattgcccagtaaggaattagtcctatccaaagctggggc 578  
 Db 718 ggcacaaagatactatctcattcattatgataaagaactcttgccatgcaaaagctgagc 777  
 QY 579 tgggcagaagaagctctgcatcaccgaacacagtcataagctttagggccagagaa 638  
 Db 778 agtcagctgaagaagctcttagcagcaatcagaactggtgctttagggagacaaagaa 837  
 QY 639 agaactcaaaagtgcttccctttaaataaagaatgctgtgttatttcccca 698  
 Db 838 agaacttga----- 846  
 QY 699 gtggtactaaagtgtgttctgtntttgtaaggtatcacagaaatcacaagaatgcaa 758  
 Db 847 -----aagttacaacaaatctagaagaagctaa 876  
 QY 759 ggaattggcataaaaggactatagctcaaaagtgctgttctgtgacttctt 818  
 Db 877 aggaattgggataaagaagctatcagcagcaaatcttatttggtgcttcttatt 936  
 QY 819 tatgaatgaacataaggaactgtctttagtgatgagaactccttgatcttaatgaga 878  
 Db 937 gatctatgacatatagtctgtgttctgtgtatggagaaactccttggttccctcagtg 996  
 QY 879 accgtgataccatcggaactgtctgtgttcttgatttcttaagttatccatagacgta 938  
 Db 997 -----atattctatggacaagaactcaacgcttcttcttctatgaattggggctt 1050  
 QY 939 ttgattggacagcagtcctcacttcttgaacacttgcgaatagccgagagctgctt 998  
 Db 1051 tagtattggacagcagctcccaagcatgttgagacttggcaacgagagagagactta 1110



```

Db      838  agaacttga----- 846
Qy      699  gfggctactaaagtgtgtctgttnttttgaaagtatacagagaatctcaagaatgcaaa 758
Db      847  -----aaggtacacaaaatttgaagaagttaa 876
Qy      759  ggaatttgcaataaaaggactatagctcaaaagtctctctgtgctgtgtactctt 818
Db      877  aggaatttggaataaaggactacacggccaaactctatgtgctgcgcttcttatt 936
Qy      819  tatgtatggaacctatgagctgtcttttgtatgtgaacctctgtatcttaatgaga 878
Db      937  gatctatgatcatatgctgtcttctgtatgtggaacctctctgtctctccagatga 996
Qy      879  acctgtatatcatcgtgagctgtctgtcttcttcttcttcttcttcttcttcttct 938
Db      997  -----atattcatgtgacaaagtactacgtctcttcttcttcttcttcttcttct 1050
Qy      939  ttgcatgtgagcagcagctccctcactcttgaaccttgaaccttgcgaatagcccgagagctgctt 998
Db      1051  tagtatgtgacagagctccccaagcatgtgaagcatgttgcgaagcagagagcagctta 1110
Qy      999  tcaatattccaggttatgtataagaaccagtatagataactttccacagcttgata 1058
Db      1111  tgaaccttcaagataattgacataataaccagcatgtgacagctatctgaaagatgtga 1170
Qy      1059  taacctgtatccatataagaagagctgtggaatttaaaagtcttcttcttcttcttct 1118
Db      1171  taacacagaataatctaaaggaaatcttgaaatcaaaaagtcttcttcttcttcttct 1230
Qy      1119  aagacatcatcaatcaatctgaagctgtgaatctcgaatgaatgaatctgtgaagacagt 1178
Db      1231  tcgaaaaagaagttaagaatcttaagggtctcaaccttgaaggttcagagtgaggagagct 1290
Qy      1179  cgctctgtctgtctcaatgtgcaatgtggagagatgaagtgatgcagcttctgcagaaggt 1238
Db      1291  ggcgctgtgttggaacagctgtgcgaggagacacagccgtgcagctgtgtgagagaagct 1350
Qy      1239  atctgtacgagatgagctgtcttctatcagtggtgagatgaagacatcagcgttaagt 1298
Db      1351  ctatgaccacacagatgagctgtctgtatctatgtgacagagacattgagccataatgt 1410
Qy      1299  ggcgcatatcagagacatattgagtggtgtatgcaagaagcctgttcttctgcggagcac 1358
Db      1411  aagcgatcttcgggaataattactcgtgtgtgtgagctcagagcctgtgttctgcaccac 1470
Qy      1359  catcagtaacataatcaagtatagcagagatgagtgatcgtatgaagaagatgagagagc 1418
Db      1471  gatagctgaaaaacatctgcgtatgcccggaaaaatgacacatgagatgagatltgagaagc 1530
Qy      1419  agcaaggaggaagcaatgctgatttattatcatagaggttctcctaataatttaacat 1478
Db      1531  tgttaaggaaagccaatgctcattatctatcatcagaactactactataatttgacacttc 1550
Qy      1479  gtaggggaaaaaaggagctcaaatgagtgagggagcagaaacagagatcgcgaattgtctcg 1538
Db      1591  ggttgagagagagagggcgccagcgtgagtggtgagcagaaagaagatcgcgcatgtccgcg 1650
Qy      1539  tgccttatctcgaaccccaagatctgattttagatgaggtcagctcgcctgtgattc 1538
Db      1651  ggccttgctgcgaaccccaagatctctctgtcgtatgagtgagcaacgtcagctcgtgacac 1710
Qy      1599  agaaagcaagatcagctgttcaagctgcagctgagagaagc 1636
Db      1711  tgaagtgaaagcagtggttccagtggtcccttgataagc 1748

```

RESULT 5  
 AAD03489  
 ID AAD03489 standard; cDNA: 4317 BP.  
 AC AAD03489;

```

XX 13-JUN-2001 (first entry)
DT
XX
XX Dog P-glycoprotein (PGP) cDNA #2.
DE
XX
XX Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
XX Canis familiaris.
OS
XX
XX Key Location/Qualifiers
FH 70..3912
FT CDS /*tag=a
FT /product="Dog P-glycoprotein (PGP) #2"
FT
XX
XX WO200123540-A2.
XX
XX PD 05-APR-2001.
XX
XX PF 28-SEP-2000; 2000WO-US26767.
XX
XX PR 28-SEP-1999; 99US-0156510.
XX
XX
XX (GENT-) GENTEST CORP.
XX
XX PI Stocker PJ, Steimel-creepi DT, Crepsi CL, Relf TC, Patten CJ;
XX
XX DR MPI: 2001-235373/24.
XX
XX DR P-PSDB: AAE00304.
XX
XX PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
XX for determining the bioavailability of drugs and for screening for dog
XX PGP inhibitors.
XX
XX PS Claim 1; Page 66-72; 11pp; English.
XX
XX CC The invention relates to dog P-glycoprotein (PGP) also referred
XX as multidrug transporter (MDR1) and nucleic acids encoding them.
XX CC The invention also includes fragments and biologically functional
XX CC variants of dog P-glycoprotein. PGP and their nucleic acids are
XX CC useful for determining the bioavailability of drugs and for
XX CC screening PGP inhibitors. They are useful for the diagnosis and
XX CC treatment of conditions characterised by PGP activity, by
XX CC reducing or increasing PGP activity in a cell. PGP nucleic acids
XX CC are used as oligonucleotide probes. Complements of PGP nucleic
XX CC acids are useful as antisense oligonucleotides, to induce a PGP
XX CC 'knockout' phenotype. They are used to prepare a non-human
XX CC transgenic animal, which are valuable as genetic models for
XX CC human diseases.
XX CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
XX CC PGP enzyme functions as an efflux pump exporting small molecules
XX CC across the cell membrane. This enzyme is a member of the ABC
XX CC transporter family.
XX
XX SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

```

Query Match 26.2%; Score 528.8; DB 22; Length 4317;  
 Best Local Similarity 62.7%; Pred. No. 4,6e-147;  
 Matches 926; Conservative 0; Mismatches 462; Indels 90; Gaps 3;

```

Qy 162 caggttaccctgattatgttgaatagtggtgctgcttcttcttcttcttcttcttct 221
Db 408 cagtgatgcctattatcagtggtgagctgtcgtgctgtgtgctgtcttaccatca 467
Qy 222 gattctcttggtatatactgacgacagacagacaaagaggttcgaaacagtttt 281
Db 468 gtttcaatctcgtgtcgtgagcagagagagacatactcaaaattgaaacatttt 527
Qy 282 tcattcagtttgacagacagacatcgctggtttgatagctgtgtacatcggtgaactaa 341
Db 528 tcattcagtttgacagacagagagatggtgtgtttgacgtgtacatgacgttggagactaa 587

```



OY	342	caaccgcatagcagcaacttgcacaaatcagtgatgtgatgagataagattgcctc	398
Dp	588	caaccgcgcatacagagacgaatgtctcccaaaatcaatgaaaggaatttgcgcgaacaagattgcat	647
OY	399	gtgtttccaacaacatgtcactcttccatcttcgatcttgccgcggagcttggttttgtaagagctg	458
Dp	648	gtctcttccaatcaatagcaacaatttccaacggtttttatagtgggggtttacaactggttg	707
OY	459	gaaaactaaccttagtgactatccatccacgctccctcttaataatgcttcacgcgcgcagcatg	518
Dp	708	gaagctaaccttggttgatttggccatbaagccctgctctcttggaattcaagccgcgacatctg	767
OY	519	ttctcagaatggttcatactcattgacccagtaagaataagtgctctattccaaagctgggc	578
Dp	768	ggcaaaagatactatctcattacttactatgataaagaactctgtgcctatgcaaaagctggagc	827
OY	579	tgtagcagaagaagtcgtctgcatcaatactcgaaacgcatagccttataggcccaagaa	638
Dp	828	agtagcaggaagaagtcctctagcaacaatcaggaactgtagattgctcttggagagacaaga	887
OY	639	agaaactccaagtgctctccctttaaataataacaagatatgcttggtttatttcccca	698
Dp	888	agaacttga-----	896
OY	699	gtggtactaagtgtgtctgttnttgyaagatatacagaactccaagatgcaaa	758
Dp	897	-----aagttacaacaanaaatttgaagaagctaa	926
OY	759	ggaatttggcataaaagaagctatagctlccaaaagtgctctcttggtgcgtgtactctt	818
Dp	927	aagaatttggatataaagaagaactatccagcgcaacatttctatgtgtgcgcttcttatt	986
OY	819	tatgatacgaacacttgaaacttgctcttcttggataggaacctcttgatcttaatgagaa	878
Dp	987	gactatacgaatcatatgctctgcttctcttgatagtggaacctccttggtctctccagga	1046
OY	879	aacctgataccaatcoggagactgctcttgctgttcttctttagtgtaatccatagacagta	938
Dp	1047	-----atatactattgaaacaggtactcaactgctctcttcttcttgattaatgggctt	1100
OY	939	ttgcatgtggagcagcagtcctcctactttgaaaccttcgcaatagcccgagcgagctgcctt	998
Dp	1101	tagatattggaagcagcatccccaagcattgaagagatttggaaacgcaagagagagcagctta	1160
OY	999	tcaatttccaaagttatgtatagaagaacccaagatagataacttccaaagcgcagata	1058
Dp	1161	tgaaatcttcaagaataatttgcacataaaccagaacatctgcagctatctcgaagatgagca	1220
OY	1059	taaacctgtaacccaatagaaaggaactgtggaatttaaaaatgttcttccaattccatc	1118
Dp	1221	taaacccagatataattatgaaggaaatttggaattcaaaaaatgttccattccagtacccttc	1280
OY	1119	aagaccatctatcaagattcttgaagaagtcctgaaatccagaatataagtcctggagaaacgt	1178
Dp	1281	tgcgaagaagaatttaagaattctttaaagggtctccaacctggaagcttcagaaatgaggaacag	1340
OY	1179	cgacttggctgcgtctcaatgttgcagtggaagtagtgaagtaagccagcttcgcaagatt	1238
Dp	1341	ggcgctggttgcggaaacagtgctgcgcgggaagagcaagcccgtagcagctgaatgcagaagct	1400
OY	1239	ataatgatacgcgaatgagtgccttatacattggtgagaaatgacatcaatagagcctttaaagt	1298
Dp	1401	ctaagaccccaacagatgcatgtgtctgatttgtagaagcaagcaatlaggacataataagt	1460
OY	1299	ggcgcatatccagacatacttgagatgtgtatgtaagaagccctgtttgttcggagcaac	1358
Dp	1461	aaggcatacttgcggaaataattctgtgtgtgagatcagagcgccgtgtgttctggccaacac	1520
OY	1359	catcagataacaatatcaagaattgagcagaatgtgtagcatgtatgaaagaaatgagaaagcc	1418
Dp	1521	gatagcctgaaaaacattctgcctatggtgcgcggaaaatgccaacatgtagagaaatttgaaagc	1580
OY	1419	agcaaaaggaaacaaatgcatgatatttatacagtagaggtcttccaataaatttaacatc	1478

Db	1501	tcgttaaggaagccatgacctgaattatcaatgaactaacctaaattgacacccct	1640
Oy	1479	ggttaagggaataaagagctcaaatatagtggaaggcagaaacagaaagatcgcaatgctcg	1538
Db	1641	ggttggaagaagagagggcccgctgtagtggtagacagaaacagaaatcgcattgctccg	1700
Oy	1539	tgccttaagtcgaacccccaagatctgattttagatgagagctacgctgccttgatc	1598
Db	1701	ggccctggtctgcacaccccaagattctctcgtgtgtagaggaacgctcagcttgacac	1760
Oy	1559	agaaagcaagtcagctgttccaagctgcacctggagaag	1636
Db	1761	tgaagtgaaagcagctggttcaggtgagccttgataag	1798
RESULT 6			
AAZ49332	standard; cDNA; 3860 BP.		
AAZ49332;	(first entry)		
14-MAR-2000	Human wild-type multidrug resistance-1 (MDR-1) cDNA.		
Human wild-type multidrug resistance-1 (MDR-1) cDNA.	Multidrug resistance; MDR-1; P-glycoprotein; transmembrane efflux pump; haematopoietic stem cell; transduction; bone marrow transplantation; chemotherapy; radiation therapy; cancer; gene therapy; gene replacement; genetic defect; thalassemia; Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion; cytokine; wild-type; ds.		
Homo sapiens.	Location/Qualifiers		
Key	1..3843		
CDS	/*tag= a /product= "Human wild-type MDR-1 protein" replace (553..555, GTT) /*tag= b /note= "cDNA sequence of G185V human mutant MDR-1 given in AAZ49333"		
MO9961589-A2.	02-DEC-1999.		
27-MAY-1999;	99WO-US11825.		
28-MAY-1998;	98US-0086988.		
(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.	Sorrentino B, Bunting K;		
MP1: 2000-072615/06.	P-PSDB: AAY58186.		
Ex vivo expansion of hematopoietic stem cells transduced with a sequence encoding human multidrug resistance-1, used for bone marrow transplantation -	Claim 10; Page 68-70; 113pp; English.		
This sequence represents cDNA encoding human wild-type multidrug resistance protein MDR-1. MDR-1 is a transmembrane efflux pump, responsible for the export of drugs from cells, particularly cancer cells. Wild-type MDR-1 shows increased resistance to etoposide and decreased resistance to vinca alkaloids compared with a mutant form (AAY58187) where the Gly at position 185 is replaced by Val. The invention relates to transducing haematopoietic stem cells with nucleic acid encoding an MDR protein			



XX 04-NOV-1999; 99US-0163508.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
PI Sornasse T, Sellhauer JJ, Watson GA;  
DR WPI; 2001-291057/30.  
XX  
XX New cell and tissue specific polynucleotides useful for diagnosis,  
PT prognosis or monitoring of treatments for disorders where the gene is  
XX associated with a cancer, immunopathology or neuropathology -  
XX  
PS Claim 1: Page 207-208; 327pp; English.

XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide  
CC sequences (I). (I) can have cytosolic, immunomodulatory and  
CC neuroprotective activities, and can be used in gene therapy. (I) and  
CC proteins (II) encoded by then are used in high throughput screening  
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,  
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or  
CC their fragments, immunoglobulins, inhibitors, drug compounds and  
CC pharmaceutical agents. Expression of (I) in a sample indicates the  
CC differentiation of embryonic stem cells into a tissue selected from  
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic  
CC tissues. (I) and (II) are used to produce an expression profile that  
CC defines a metabolic or developmental process, treatment, condition,  
CC disease or disorder. The gene profile can be used for diagnosis,  
CC prognosis or monitoring of treatments and for investigating a  
CC predisposition to a disorder where the gene is associated with a  
CC cancer, immunopathology or neuropathology.  
XX

Sequence 4349 BP: 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Query Match 25.4%; Score 512.8; DB 22; Length 4349;  
Best Local Similarity 62.0%; Pred. No. 2.7e-142;  
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;

QY 162 cagggtgaccggtatattggtggaatagtggtctgctgatttgggtactataa 221  
DB 461 caggatagctctatttaagtggaattggtggtggtggtggtggtggtggttca 520  
QY 222 gatttcctggtgattatactgacgacgacgacgacgacgacgacgacgacgac 281  
DB 521 gtttcatttgggtcctgctgacgctggaagacaataacaaatagaacagtttt 580  
QY 282 tcatcagtttggacagacatcgctggttgaatagctgtgacatcggtgaactaa 341  
DB 581 tcatgctataatgacgacgagagatagctggttgaatgacagatgtgggagctaa 640  
QY 342 cactcgatgaagacatgac---aaatcagtgatgtgattggaatagatgtcct 398  
DB 641 caccgacttaacagatgagctcctaagatlaagatgattgtgcaaaattggaat 700  
QY 399 gttgttcaaacagtgactcttctgattggtcctgacgagttgtgttgtaaggagct 458  
DB 701 gttcttcagtaatggaacatcttcaacgtgtttatagtaagattcaacgtgtgtg 760  
QY 459 aaactacaccatgactctatacagctctccttataatggtctcagcgacgacatg 518  
DB 761 gaagtaacccttgattgttgcatcagctcgttcttgtagctgtgacgtcgtgtcg 820  
QY 519 ttcttgatggtcatctatctatgacagtaagaatagtgctatccaagctggggc 578  
DB 821 ggaacagatacatcttcaattacgtaaaagaaacttaagacttaagcaaaactggaac 880  
QY 579 tggcgagaagaagctctgtcaatccgaacagtcacatagcttagggccagagaa 638  
DB 881 agtagctgaagaggtctgtgacgacatagaactgtgattggttgaggcgaaagaa 940  
QY 639 agaactcaaaagctcttcctttaataataacaagatagtctgtttattttcccca 698  
|||||||

DB 941 agaactga----- 949  
QY 699 gtagctactaaattgtgtcttnttttaaggtatacagcaatcaaatgacaa 758  
DB 950 -----aagttacaacaaattagaagactaa 979  
QY 759 gatttggcataaaaggactatagctcaaaagtctctgtggtgctgaactct 818  
DB 980 aagaattggataaagaagcattacagccaatattctatagtggtcgttctcgt 1039  
QY 819 tatgaatgaacctatgactgcttcttggataggaacctcctgattctaatgaga 878  
DB 1040 gatcatgactctatgctcctgctgctgattgagggacacctgctcctcagggga 1099  
QY 879 acctgatatccatccgactgtcttctgttcttctttagtgaatccatagcaatga 938  
DB 1100 -----atatctatggaacagactcactatcttcttctgtatttaattggtgctt 1153  
QY 939 ttgcatggagcagcagctccctcaacttgaacctcgcaatagcccgagagctcct 998  
DB 1154 tagtgttgacagcagctcccaagcatggaacattgcaaatgcaagagagcagctta 1213  
QY 999 tcatatttccaggtattgataagaaccagatagataaactttccacagctggata 1058  
DB 1214 tgaattctcaagataattgataatagaacagatattgacagctatgcaagatggca 1273  
QY 1059 taacctgataccaagaagaagcagctggaatttaaaatgttcttcaattccarc 1118  
DB 1274 caaacagataatlaataaggaatttgaattcagaatgtctcctcattgataccac 1333  
QY 1119 aagaccatcatcaagaattctgaagctgaatccagaatlaagctcgagagacagt 1178  
DB 1334 tcgaaaagaagtgatagattcttgaagggcctgtaacctgaagtgccagagtgagcag 1393  
QY 1179 cgcttgtcgtggtcctcaatgagcagtggaagagtagctgagctcctgcaaggtt 1238  
DB 1394 ggcctgttgtaaaacagcagctgtggaagagcaacaacagctgataagaggtc 1453  
QY 1239 atatgacccgagatgagcttcatcaatggttgatgagaatgacatcaaggttaagt 1298  
DB 1454 ctatgaccccaagagggatgctcagtggttgaatgacagagatattagaccataatgt 1513  
QY 1299 ggcgcatctcgagacatattggaattggttagtcaagagcgtgtttctgggacac 1358  
DB 1514 aaggttctcaagggaaatcatctgtgtgtgagtcaggaacctgtatgtttgcccacc 1573  
QY 1359 catcagtaacaatalcaatgaatgagcagatgagtgtgactga tgaagatgagagagc 1418  
DB 1574 gatagctgaaaacatctgctatggtcgtggaatgtcaccatgagatgagattggaagc 1633  
QY 1419 agcaagggaagcaaatgctatgatttatacatgagattcctataaattataacatt 1478  
DB 1634 tgcagaaggaagcaatgctatgacttatacgtgaatcagctcctataattggaaccc 1693  
QY 1479 ggttaggggaaaagagcagcaatgagtgagtgagggcgaagaaagagatcgaaattgctg 1538  
DB 1694 ggttggagagagagggcccaatgagtgagtgagggcgaagaaagagatcgaaattgctg 1753  
QY 1539 tgccttagtgcgaaccccaagattctgatttaagatgagctgagctgctgattc 1598  
DB 1754 tgcctgtgtgcgaaccccaagattcctctgtgtgagggcgaagcttggaacac 1813  
QY 1599 aagaagcaagtcagctgttcaagctgctgactggaag 1636  
DB 1814 agaaagcgaagcagtggttcaagtggtccttgagataag 1851  
|||||||

RESULT 8  
AAQ72872  
ID AAQ72872 standard; cDNA; 4646 BP.  
XX  
AC AAQ72872;  
XX

DT	21-JUN-1995	(first entry)
DE	Human multidrug resistance gene (MDR-1) encoding a p-glycoprotein.	
XX		
XX	Human: tyrosinase; compound: hair; follicle; liposome; lipophilic;	
KW	lipophobic; stratum corneum; melanin; hair dye; glycoprotein; alopecia;	
KW	chemotherapy; transforming growth factor; growth stimulant; aromatase;	
KW	cyclosporin A; macromolecule; polymer; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	425..4267
FT		/*tag= a
FT		/product= p.glycoprotein
XX		
PN	W09422468-A.	
XX		
PD	13-OCT-1994.	
XX		
PF	01-APR-1994;	94WO-US03634.
XX		
PR	02-APR-1993;	93US-0041553.
XX		
PR	13-JAN-1994;	94US-0181471.
XX		
PA	(ANT1-) ANTICANCER INC.	
XX		
PI	L1 L, L1shko VK;	
XX		
DR	WPI: 1994-332816/41.	
DR	P-PSDB; AAQ72872.	
PT	Liposomes for delivering protein, nucleic acid etc. to hair	
PT	follicles - e.g. to restore hair colour, prevent hair loss during	
XX	chemotherapy, stimulate hair growth etc.	
XX		
PS	Claim 8; Page 70-77; 100pp; English.	
XX		
CC	The nucleotide sequence of the human multidrug resistance (MDR-1) gene	
CC	which encodes a p-glycoprotein. This is an example of a compound which	
CC	can be delivered to hair follicles via a novel liposome composition.	
CC	The liposomes are comprised of a lipophilic or lipophobic compound which	
CC	will selectively target the hair follicle (via the stratum corneum)	
CC	without damaging or unwanted effects on cells outside the follicle.	
CC	Compounds e.g. p-glycoprotein, can be delivered to the hair follicle for	
CC	treatment of chemotherapy-induced alopecia. Other compounds targeted at	
CC	hair follicles can include: tyrosinase (or the DNA encoding it -	
CC	AAQ72871), melanin or hair dyes (to restore hair colour or condition);	
CC	human transforming growth factor-alpha (AAQ72873)(for reversal of wavy	
CC	hair); hair growth stimulants (e.g. cyclosporin A or aromatase) or	
CC	antitense sequences. The method allows compounds (e.g. macromolecules	
CC	or polymers), which would not normally reach the hair follicles, to be	
CC	delivered to these target areas.	
XX		
SO	Sequence 4646 BP; 1371 A; 893 C; 1128 G; 1254 T; 0 other;	

Query Match	25.4%;	Score 512.8;	DB 15;	Length 4646;
Best Local Similarity	62.0%;	Pred. NO. 2.8e-142;		
Matches 916; Conservative	0;	Mismatches 472;	Indels 90;	Gaps 3;

OY	162	caggatggccctgtrattgttgggaatagagtgatgcgcgccttgaatttttggtatcaata	221
Db	760	caggatggccctatattcatcagatgaaatgtgctgggggtgcctgtgtcgtcttaattca	819
OY	222	gatttccttfggattatataactgcagcagcagacgaaccaagaagattcgaacaacgtttt	281
Db	820	gatttcattcttggcgccgagcctggaagagaacaatacacacaaatttgaacaacggtttt	879
OY	282	tcatcagattttggacagagacatcgcgtgtgttttgatagctgtgtacatcgtgtgaactaa	341
Db	880	tcatcctataagcagcacgagagatagctgtgttttgatgtgcaacatgtttggagacctaa	939

QY	342	cactcgcgcatgcagacacttgaac-----aaatcagctgatacgtgatctggaataagtaattgctct	398
Db	940	caccgcgacttccagatacgtatcttccctaagaattcaatgaaatttgcttggcacaataattgaaat	999
QY	399	gtctttccaanaacgtgcactcttccattccagatttgcgcgcgagcttgcttggtggaagcgtc	458
Db	1000	gtcttccagtcgaatgvcgaacatttcttcaactcgttggttataagtaagatttacaacgttgctc	1059
QY	459	gaaacacccttgaatgcataccatccacgctccctcttataatgcttcaagcgcgacacgt	518
Db	1060	gaagctaacctcttgatcttcttgccatccagctccctcgtcttcttgacgtccagctgctgctc	1119
QY	519	ttctagaatggtcattcccatcttgcacccagtaagaataagtcctattccaaagctgggac	578
Db	1120	ggcaaaagctactatcttcatcttacttcaataaagacctcttaggtatgtaacaaagctggac	1179
QY	579	tgctgcagaaagaatctctgtccatcaatccgaacgctacagctcttaaggcccgagaa	638
Db	1180	agtaactgaaagaggtctcttgccagcaattagaaactgtgattgattcttggagacaanaaa	1239
QY	639	agaattccaagaagcttccctctttaaataacaagatacgtcttgcttatttcccca	698
Db	1240	agaaacttga-----	1248
QY	699	gtggctactaagctgtctgtctgtnttctgtaagctatacacagaactccaagaatgycaaa	758
Db	1249	-----aagttacacaaaataattagaagaagctaa	1278
QY	759	ggattttgacataaagaagcattagcttcaaaaagctctctcttgctgctggttactctt	818
Db	1279	aagaatttggataaagaagaactatccacgcaaatcttataagtgctgcttctcctgt	1338
QY	819	tatgaatgaaacctatgactctcttcttgatagtcgaaacctcgtatcttcaatgcgaga	878
Db	1339	gatactatgcacttaatgctctgctgcctctcgtatgtagaacactggtccctccaaggga	1398
QY	879	accttgcataaccatcgcggaactgctcttgctgttcttctttagtgtaatccatagacgtta	938
Db	1399	-----atatcttatttgacaagaactcgaactgattcttcttgattataatgryggctt	1452
QY	939	ttgctattggacagcagctccctccacttggaaacctcggaaatagccgaggggcgtct	998
Db	1453	tagtgttggacaagcactcccaagcattggaagcatttgcanaagcagaagggacagctta	1512
QY	999	tcatatttccaggcttatgataaagaaccagctatagataacttccacagcgcctgata	1058
Db	1513	tgaaacttccaagataatttgataaagccaaagtattgacgctatctcgaaagatgryggca	1572
QY	1059	taaacctgtaaccatagaaaggaactgtgtgaatttaanaatgttcttccaattccact	1118
Db	1573	caaacagataattttagaagaaatttggaattcagaatgtccacttccatgtaaccact	1632
QY	1119	aagaccatctatcaagattcttgaaggtctcgaattccgaattaaagtctcgagagacagt	1178
Db	1633	tcgaaaagaagttaagatcttgaagggcctcgaaccctggaagtgtcagagtgtggcagaagct	1692
QY	1179	ggccttcttgcttccatagtcgacgttggaaagtaagctagctccagcgtctcgaaggtt	1238
Db	1693	ggccctgtgttggaaaacagtggtcgttggaaagacacaacagctccagctgatacgaaggt	1752
QY	1239	atatatcccgatgatagtgcttattcaatgcgtggaatggaatlgacatccagaagctttaaagt	1298
Db	1753	ctatgaccccaacagagggaatggtgcagtgttctgtagacaggtatattagaaccataatgct	1812
QY	1299	gcgcgcatcatccagacatattggaatgtgtctagtcgaagcctgtcttcttcgggaccaa	1358
Db	1813	aaggttcttccaggaataacatctgtgtcgtgaaatccaggaacctgtattcttggccaacac	1872
QY	1359	catcagtaacaattccaagtatgacagagatgtgtgacatgacgtatgaagaatggaagagc	1418
Db	1873	gatagctgaaaaacattcgctatggtcgttgaanaatgtccacaaatgataagatgagaagaac	1932
QY	1419	acccaagggaaacaaatcgtatgatttataccatggaagtttccataaatttatacat	1478

Db	1933	ttgcaaggaaacccatgctctatgcttattcaagaaactgctccataaatltyacaccc	199
QY	1479	gttaggggaaaaaagaagcctcaaatgagttggaaggcagaanaacagagatcgcattgctcg	153
Db	1993	ggttcggagagagagaggcccaagcttgagttggtggcagaagaacagagatcgcattgcaag	205
OY	1539	tgacctagttcgaaacccccaagattctgatttagatgaggaagtcagtcgtccctgatttc	159
Db	2053	tgccctggttcgcacaccccaagaattccctcctgctgtagagcgacagcagccttgacac	211
OY	1599	agaaagcaagtcagctggttcaagctgacgtcggagaag	1636
Db	2113	agaaagcgaagcagtggttcaagtcagtcgtcgtgataaag	2150
RESULT 9			
ID	AA294738	standard; cDNA: 4646 BP.	
XX	AA294738;		
XX	01-AUG-2000	(first entry)	
XX			
DE	Human ATP binding cassette ABCB1 (MDR1) cDNA.		
KX	ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;		
KW	atherosclerosis; lipid disorder; dyslipidemia; psoriasis;		
KW	lupus erythematosus; diagnosis; gene therapy; MDR1;		
KW	multidrug resistance; chromosome 7q21; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200018912-A2.		
PD	06-APR-2000.		
XX			
PF	21-SEP-1999; 99WO-EP06991.		
PR	25-SEP-1998; 98US-0101706.		
XX			
PA	(FARB ) BAYER AG.		
XX			
PI	Schmitz G, Klucken J;		
XX			
DR	WPI: 2000-293151/25.		
XX			
PT	Adenosine triphosphate binding proteins useful for identifying agents		
PT	for treating atherosclerosis and other inflammatory disorders -		
XX			
PS	Claim 9; Page 110-112; 154pp; English.		
XX			
CC	The present sequence is that of human ATP binding cassette		
CC	subfamily B protein ABCB1 cDNA. The cDNA was identified using a		
CC	differential display method in which monocytes from peripheral		
CC	blood were subjected to macrophage differentiation and cholesterol		
CC	loading with acetylated low density lipoproteins and subsequent		
CC	degrading with high density lipoprotein (HDL3) to identify		
CC	cholesterol sensitive genes. The gene maps to chromosome 7q21		
CC	and is also termed MDR1 (multidrug resistance). The invention		
CC	provides cholesterol-sensitive ABC genes (see AA294734-63). These		
CC	genes, and polypeptides encoded by them, can be used for diagnostic		
CC	and therapeutic applications, and for biochemical or cell-based		
CC	assays to screen for pharmacologically active modulator compounds		
CC	useful for the treatment of lipid disorders, atherosclerosis or		
CC	other inflammatory diseases such as psoriasis and lupus		
CC	erythematosus.		
XX			
SO	Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;		
Query Match 25.4%; Score 512.8; DB 21; Length 4646;			
Best Local Similarity 62.0%; Pred. NO. 2.8e-142;			

Matches	915: Conservative	0: Mismatches	472: Indels	90: Gaps
QY	162	caggttgaccctgataatgattggaatagtggtgtgccttgatcttgggtacataca	221	
Db	760	caggtatgacctatattacagtagtaattgtgtgggggtcgtgtgtctgtacatca	819	
QY	222	gatttctctgtgattatatactctgacacgacagacaaagagatctgaaaaagtttt	281	
Db	820	ggtttcatcttgtgacctgagcctggaagacaaatacaaaaatttgaataacagtttt	879	
QY	282	tcattcagttttgacagagacacgcgcgttttgaatagctgtgacacgggtgactaa	341	
Db	880	tcaagtcataatgtgcgaagaagatagaccggtttgtgtgtgcagatgtcttggggactta	939	
QY	342	cactgcgatcacagacattgac---aaatcagatgatagtatctggaataaagatgtcct	398	
Db	940	caccgcgacttacagatgattgtctctaaagtttaatgagttattgttggacaatatgtgat	999	
QY	399	gtgttttccaaaacatgtctactctttctcgatttggcctgcgcagttgttgggtgaaggctg	458	
Db	1000	gtctcttcaagtcataatggacaattttccatcttggtttatagtaggatttcaacgttgctg	1059	
QY	459	gaacatcaccttagtgactatccatcacgcgtccctcttaaatgctgtcagcggcagctg	518	
Db	1060	gaagctaacccctgtgtgatttggcctatcagctccctgtcttcttgacgtcagctgcgtctg	1119	
QY	519	tcttagagatgtctactctacttgacccgaataagatgaatgtccattcccaagctgaggc	578	
Db	1120	ggcaagaatactatcttcatcttactatgataaagactcttagtgatgataaaggctgggc	1179	
QY	579	tgtgacgaagaagaatctctgtcatalcaatccgaagaatgcataagctttagggcccgagaa	638	
Db	1180	agtagctgaagaaggtctctgtgcacgaatctgaactgtgattcttggaaagcaaaagaa	1239	
QY	639	agaacttcaaaagtccttccctttaaataataacaagatagtgtgtttatcttcccca	698	
Db	1240	agaacttga-----	1248	
QY	699	gtgctactaagatgtgtgttnttggtaaggtatacacagaatctccaagaatgacaa	758	
Db	1249	-----aagatacaacaanaatttagaagaagctaa	1278	
QY	759	ggatttggcataaaaggactatagcttcaaaagtgctctctgtgtgtgtgtactctt	818	
Db	1279	aagaattgggataaagaagactatcacgcaaatcttctataagtgctgtccttccgt	1338	
QY	819	tatgaatggaaacctatgactgtcttttggataggaacctctgattcttaagggaa	878	
Db	1339	gatactatgacatctatgtccttgcgtctctgtgataaggacacctgtgctctccagggga	1398	
QY	879	acctgtgatatccactcgagcagctgtctctgtgtgttcttctttagtgaatccatagcagta	938	
Db	1399	-----atatctatttggacaagaagctacactgtatctcttctctgtattaatggggctt	1452	
QY	939	tgtcatttggagcagcagctccctcaacttggaaaccttcgcaatagcccgagagctgtcct	998	
Db	1453	tagtgttggacagcactcccaagcatttgaagcatttggaaatggcaaggagcagctta	1512	
QY	999	tcatatttccagtttatgtatagaacaccgataagataaactttccacgcgcgtgata	1058	
Db	1513	tgaacttctcaagaatatgtatataaagccaagatctatgacgctatctcgaagatgggca	1572	
QY	1059	taaacctgaatccatagaagaagaactgtggaatttaaaaatgttcttcaatatcacac	1118	
Db	1573	caaacagataatcttaagggaatttgggaattccgaaatgttccacttccagttaccac	1632	
QY	1119	aagacactctatccaagatcttgcgaaggctgaatctcgaatgaagctctggaagacagt	1178	
Db	1633	tcgaaaagaagttaaggtcttggaaaggccctgcacactggaagtgacagatgggcagaga	1692	
QY	1179	cgctctgtgcgtctcaatgcagctgggaagataagctafagcccgctcttgcagaagct	1238	
Db	1693	ggccctgttctggaacacaggtgtctgttggaaagcacaacagttccagctgtgtgtcaga	1752	



```

Db 1339 gatctatgatcttatgtctgtgctctgtgatggaaccactgtgtcctctcaaggga 1398
OY 879 accitgatataccatcgagcagctgtctgttcttcttaagtatccatagacgtta 938
Db 1399 -----atactcatgtgacaagtaactcactgtatcttcttctgtatattgtgggctt 1452
OY 939 ttcgacttggagcagcagctcctcacttttgaaccttcgcaatgcccggagagcctt 998
Db 1453 tagtcttggaacagcgtcctccaaagcatgtgaacatttgcaaatgcaagggcgacctta 1512
OY 999 tcatatttcagagttatgtatagaagaaccagtatagataactttccacagcttgata 1058
Db 1513 tgaactcttcaagaatgaatgaataaagccaagtatgtgcagcatcttccaagagtgggca 1572
OY 1059 taacctgataccatagagaagcagctgtggaatttaaaatgttcttcttaattaccac 1118
Db 1573 caaacccagataatgaagaagaatttgaaatccagaatgttccactgaattaccac 1632
OY 1119 aagaccatctatcaagatctgaagaagtcgaatctcagaatgaagtcggaagacagt 1178
Db 1633 tcgaaaagaagcttaagatcttgaaggcctgaacctgaagcgagcagagtgagcagacagt 1692
OY 1179 cgctctgtctggtctccaatgagcagtggaagagtcagtgatgcagctctcgaagagtt 1238
Db 1693 ggcctctgttgaagaacagctgtgctgtggaagagcacaacagcagctgatgcaagagct 1752
OY 1239 atatgaccgagatgtgcttatactatgtgtgatagaatgacatcagaagctttaatt 1298
Db 1753 ctatgacccacagagggagtgatgctgtgtgacagagatataagacataaattgt 1812
OY 1299 ggcgacatcaacgacacatctgagtgagtgatgcaagagcctgtttgttcggagacac 1358
Db 1813 aaggttctcaacggagaaatcatctgtgtgtgtgagtcaggaacctgtattgttgcacac 1872
OY 1358 catcgatacaatatacaagtatgagcagagatgctgtacgtatgaaagatggaagagc 1418
Db 1873 gatactgtaagaacatctgcgtatgcgcgtgaatgttcacacatgagatgagaaagc 1932
OY 1419 agcaagggagagaatgctgtatgatttctatcagagttcctctataaatttaatacatt 1478
Db 1933 tgcagaaggaagcactgtcctatgacttctatgaactgcactcaaatltgacacccct 1992
OY 1479 ggtaggggaagaaagagcgaatagtgagggcagaagaagatcgacatgtctcg 1538
Db 1993 ggttgagagagagaggggcccagtgagtggtgtggcagaagcagaagatcgacatgtcag 2052
OY 1539 tgccttgaatcgaaaccccaagatctctgattttagatgagagtcagctgcgcctgagatc 1598
Db 2053 tgcctgtgttcgcaaccccaagatcctcctgtctgagatgagggccacgttcagccttggacac 2112
OY 1599 agaagcgaagtcagctgttccaaagctgcagctgagaag 1636
Db 2113 agaagcgaagcagtgctgagtgagtgcttgataag 2150

```

RESULT 11  
AAT13394  
ID AAT13394 standard; DNA; 6505 BP.

AC AAT13394;  
XX 24-JUN-1996 (first entry)  
XX Hybrid vector pSF-MDR.  
DE Hybrid; vector; gene transfer; gene therapy; haematopoietic stem cell;  
XX Hybrid; vector; gene transfer; gene therapy; haematopoietic stem cell;  
KM retroviral; murine embryonic stem cell virus; MESV;  
KM Mooney murine sarcoma virus; (MOMUSV);  
KM Friend murine leukaemia virus; F-MuLV; ds.  
OS Synthetic.  
XX DE19503952-A1.

```

XX 14-MAR-1996.
PD 07-FEB-1995; 95DE-1003952.
XX 08-SEP-1994; 94DE-1431973.
XX (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX Baum C, Osterlag W, Stocking-harbers C, Stockingharbers C;
XX WPI; 1996-152306/16.
DR Hybrid retroviral vectors - for gene transfer into haematopoietic
XX stem cells
XX
XX Disclosure; Page 25-29; 42pp; German.
XX
XX New hybrid vectors comprise (1) a leader region including the U5
XX region and trna primer binding site of murine embryonic stem cell
XX virus (MESV) or Mooney murine sarcoma virus (MOMUSV), and (2) a 3'-
XX LTR including the U3 and R regions of a Friend murine leukaemia
XX virus (F-MuLV). The vectors are useful for ex-vivo or in-vivo gene
XX therapy. High levels of gene transfer can be achieved in
XX haematopoietic stem cells and their myeloid (non-lymphatic) progeny.
XX pSF1, pSF2, pSF3 and pMW1 (sequences given in AAT13390-T13393) are
XX examples of such vectors.
XX Vector pSF-MDR (sequence given in AAT13394) is based on the
XX MESV vector R224.
SO Sequence 6505 BP; 1719 A; 1506 C; 1606 G; 1674 T; 0 other;

```

Query Match 25.4%; Score 512.8; DB 17; Length 6505;  
Best Local Similarity 62.0%; Pred. No. 3; 5e-142;  
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;

```

OY 162 caggttgacccgttatatgcttggaatagtggtgtcgtcctgtatattgtgttactata 221
Db 2152 caggtatgctctattatgaagtggaattgtgtcgtggtgtcgtcgtcgtcgtcgtcgtc 2211
OY 222 gatctcctgtgtatataactgcagcagacagacagagagatcgaaacagtttt 281
Db 2212 ggtttcatttggctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2271
OY 282 tcatcagtttggcacaagacatcgctgtgttgaatgctgtgacatcggtgaactaa 341
Db 2272 tcatgctataatgcagacagagatagctgtgttgaatgctgtgacagatgttggaactaa 2331
OY 342 cactcgcatgacagacatgtac-----aaatcagtgatgtgtatggaataagattgtctc 398
Db 2332 caccgacttaacagatgagtgctcctaagatgaatgaatgttggaataaatttgaat 2391
OY 399 gttgttcaaaacatgctactcttctgattgagtcggcggcagttgtgtgtgaagggctc 458
Db 2392 gttcttccatgaatgagcaacatttctcctgtgttataagatagataacacgtgttg 2451
OY 459 gaaactaccctagtgactctatccacgtctcctcttataatgctgtccagcgacagatg 518
Db 2452 gaagctaaccttgattgttgccatcagtcagtcctgtcttcttgacgtcagcgtctgtc 2511
OY 519 ttctgagatgcatctcatcttgaaccagtaagaaatgaatgaatgctctatccaagctggggc 578
Db 2512 ggcacaagatactcttcttcttcttgaataaagacacttgaatgcatgaagaagctggagc 2571
OY 579 tgtgcaagaagaagcttgcatactccagaacagatcagatctatagggccagagaaga 638
Db 2572 agtagctgaagaggtcttgcagcaatagaactgtgattgtgagtgagggcaagaaga 2631
OY 639 agaactccaaggtcttctctttaaataaacaagatagctgtgttatttcccca 698
Db 2632 agaacttga----- 2640

```







Best Local Similarity 61.9%; Pred. No. 7.7e-142;  
Matches 915; Conservative 0; Mismatches 473; Indels 90; Gaps 3;

```

OY 163 caggttgaccgcttattatgttggaaatagttgttgccttgaattttgtttactaca 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 cagatatgcttattataagtcgaatttgcctgggtgcctggtgtgttactacaa 395
OY 222 gatttcctgttgatataactgcagacagacagacaaagagatcgaanaacagtttt 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 ggtttcattttgtgtcctgcagctgcagtcgaagacaataacaaattagaaacagtttt 455
OY 282 tcaatcagtttggcacaagacatcgctggtttgtagctgtgacatcogtgaacttaa 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 456 tcatgataatagtcgacagagataagctggtttgtagtcgaagatgttggagagcttaa 515
OY 342 cactgcagacagacatgac---aaatcagtagtcgttggatggagaatgtctct 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 caccgcacttaacagatgatacttaagataatgaagtattatgtgcacaaatttgaat 575
OY 399 gttgttcaaaaacatgtctacttttcgaattgcctgcagctgtgttggtagaaggctg 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 gttcttcagtcacatgacacattttcactggtttatagtagattacacgtgttg 635
OY 459 gaaactcaacctagtaactctatccacgtctcctctataatggtctcagcgagcatg 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 gaagctaacctgtgatttgcctcagctcagctgtcttgcgaactgcagctgtctc 695
OY 519 ttcttagtagtgcattcatcttgcacgtaagataagtaagtcctatccacgtggcg 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 ggcacaagaatactctctacttactctgaataagaactcttgcgtatgcacaaagcttgc 755
OY 579 tgtgcagaagaagctctgtcatcaatccgcgaagctacatagcctttagggccagagaa 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 756 agtagctgaagagtgctctgcagcaattagacgtgattgcatltagagacaaagaa 815
OY 639 agaacctcaaaaggtcttccctttaaataacagaatagctgtgttatttcccca 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 816 agaacctga-----aaagtaacacaaaatttagaagaagctaa 854
OY 699 gtgtgcataagtgtgtctgtnttgytaaggtatcacagaaatcacaagaatgcacaa 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 -----aaagtaacacaaaatttagaagaagctaa 854
OY 759 ggaatttgcaataaaagaactatagcttcaaaagtctctgtgtgtactctt 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 aagaattgggataaagaagctatcacgcaatattctatagtgctgtcttccgct 914
OY 819 tatgaatggaacctatggaactgtcttggatgtagaactccttgattcttaattgaga 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 915 gatctatgcatcttatgtcctgcctctgtgtatggagaccccttgcctcagagga 974
OY 879 accctggatataccatcgagactgtctgtgttcttcttgaatgtaacatagcagtta 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 975 -----atatcttctgacaagaactacactgaatcttcttctgataatagggcctt 1028
OY 939 ttgcattggagcagcagcttccctcacttgaacacttcgcaatacgcagagagctgcct 998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1029 tagtgctggacagcagctcctcaagcatgttagacattgcaaatgcaagagagcagctta 1088
OY 999 tcatatttccaaagttatgtaagaacccagatataactttccacagctgtgata 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1089 tgaactctcaagataattgataaagccaagatltgacagctatttcgaagagtgagca 1148
OY 1059 taacactgaatccatagagaagaactgtgaaatttaaaaatttcttctcaattcactc 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1149 caaacgcgataataatlaagggaatctggaatccagaatctcaactcagttaccatc 1208
OY 1119 aagacatctatacaagatctcgaagagcttgaatctcagaatlaaagcttggagagacat 1178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1209 tcgaaagaagtttaagatcttgaagggccttgaactgaagtgacagagtggcagacaggt 1268
OY 1179 cgccttgctggctcaatgagcagtggaagagtaagctagtcagcttctgcagaggtt 1238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1269 ggcctgttggaacaagctgctgtgggaagagcacacaagctcagctgtgacagagct 1328
OY 1239 atatgacccgagatgctgtcttcatgctgtgagtagaagatcagaatcagcttaagt 1298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1329 ctatgacccacagagagagtagtgcagttgtgtagacaggaatattagacataatgt 1388
OY 1299 ggcgccttaccgagacatttggagtgtagtcagaagacccgttctgttgcggacac 1358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1389 aaggttctacagggaaatcatctgtgtgtgtagtcaggaacccgtattgttgcacac 1448
OY 1359 calcagtaacaataataagatagtagcagatgtagtgcagtagaagaatgtagagagc 1418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1449 gatagctgaanaacatttcgcatagcgttgaanaatgcatcattgatatgtagaagac 1508
OY 1419 agcaagggaagcaaatgctgataatttcatcagagatttccataatattataat 1478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1509 tgcacaagaaagcaatgccttacttactatgaatgaactgcctcactaataattgacacct 1568
OY 1479 ggttaggggaaaaaggaactaaatagtagtgagggcagaacacagagatgcgaattgtcg 1538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1569 ggttgagagagaggggcccagttgagctgtaggcagaagacagagagatgcgcatltagc 1628
OY 1539 tgccttagtcgaaccccaagatctgattttagatgaggtacagctgcctgagatc 1598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1629 tgcctgtgtgcgaaccccaagatcctcctgcgtgtagtaggccaagctcagcttgcagac 1688
OY 1599 agaaagcaagtcagctgttcaagctgcagctgagagag 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 agaaagcgaagcagtggttcagctgctgtgataaagg 1726

```

```

RESULT 14
AAN70752 standard; cDNA; 4669 BP.
ID AAN70752
XX
AC AAN70752:
DT 21-MAY-1991 (first entry)
XX
XX Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones
DE lambda-HDR10.5 and 104.
XX
XX Chemo-therapy resistant tumour cell; P-glycoprotein; ss.
XX
XX Homo sapiens.
XX
XX OS
XX FH
XX Key Location/Qualifiers
FT CDS 425..4267
    /tag= a
XX
XX W08705943-A.
XX
XX PD
XX PD 08-OCT-1987.
XX
XX PF 26-MAR-1987; 87WO-US00758.
XX
XX PR 01-AUG-1986; 86US-0892575.
XX PR 28-MAR-1986; 86US-0845610.
XX
XX (UNII ) UNIV OF ILLINOIS.
XX
XX PA
XX PL Ronlison IB, Pastan IH, Gottesman MM;
XX DR
XX DR WPI; 1987-291656/41.
XX
XX P-PSDB; AAP70452.
XX
XX DNA for multi-drug resistance in human cells - used to detect
XX chemotherapy-resistant tumour cells and for producing
XX polypeptide(s) for diagnosis and therapy
XX
XX Claim 2(a); Table 5, pp30-39; 61pp; English.
XX
XX The human multi-drug resistant KB carcinoma cell lines were used as
CC

```

CC the source of the mdrl gene nucleic acid sequences (AAN70751). To  
CC obtain cDNA clones of the mdrl gene (AAN70752), poly (A) and RNA was  
CC used. Analysis of the Aa sequence presented in (AAN70752) indicates  
CC that the mdrl gene product is likely to be a transmembrane protein.  
CC The presence of transmembrane domains and potential glycosylation  
CC sites is consistent with the mdrl protein being related to the  
CC p-glycoprotein.

50 Sequence 4669 BP; 1393 A; 896 C; 1126 G; 1254 T; 0 other;

Query Match	25.38	Score 511.2	DB 8	Length 4669
Best Local Similarity	61.98	Pred. No. 8.6e-142		
Matches 915; Conservative	0	Mismatches 473;	Indels 90;	Gaps 3

OY	162	caggttgacccctgattattatgttggatagtgctgtgcttgcctgatttttggttaataca	221
Db	760	caggttgacccctgattattatgacgtggaaatgtgtgctgggtgtgctgtgtgtcttaacatca	819
OY	222	gatttcccttggtgttataactgcgaacacacagcagcagaagatctgcgaacacattttt	281
Db	820	ggttcaattttgtgtccttgccagctgtgaagacaacaataacacaatttagaaaaagttttt	879
OY	282	tcattcagtttttggcagacagacatcgcgctgtgttgaatgacgtgtgaacatcgtyaacctaa	341
Db	880	tcattcagctataatgagcagaagagatagctgtgttttgaatgtgcacagatgttggtgaactaa	939
OY	342	caccctgcatacgacgaccttgc---aaatcagtgtgtgtattgttgagataaattgtcct	398
Db	940	caccctgccttaccagatgatgtcctctaagataaatacagttatttgttgaaanaatttggat	999
OY	399	gtgtttccaanaacatgtgtcacttttccagatttgcagtcggcgagctgtgtttgtgaagggct	458
Db	1000	gttcttcagtgcaatgtgcacacttttccactcgtgttttaagttagatttaccgctgtgt	1055
OY	459	gaacacacccctagtgacgtctatccacgctcctccttaaatgtctcagcggcagacatg	518
Db	1060	gaagctaaccccttggatttttggccatcagtcgctcgttcttcttgagctgcagctgcgtctg	1119
OY	519	ttcttagatgttcattccctcatgtgcccagtgaaagatataagtgctctatccaaagctgggc	578
Db	1120	ggcagaagatctatctcattcttacttgaataaagacactcttagcgtatgacaaagctcggagc	1179
OY	579	tgtagcgaagaagaagctctgtccatccatccgaacagcatcagctcttaaggcccaaggagaa	638
Db	1180	agtcagcgaagagagctctctgtgcagcaatctgaactgtgatcttcttggaggaacaaagaa	12339
OY	639	agaaccttcaaaagtgcttcttctttaaataaacaagatatagtctgtgtatttctccca	698
Db	1240	agaacttga-----	1248
OY	699	gtgtgctactaagtgtgttctgttnttttgaagatatacaacgaatctcaagaatgcaaa	758
Db	1249	-----aaggttacaacaaaatttgaagaagcctaa	1278
OY	759	ggaatttggcataaaagaagcctatagcttcaaaagtgctctgtgtgcgtgtactctt	818
Db	1279	aagatctgggataaagaagcctattacagcccaattattctatagtgctgtccttccgtc	1338
OY	819	taagatgtgaacatcagtgactgtccttttggatataggaaacctctgatttcttaatgaa	878
Db	1339	gactctatgcatctatgtcctgtgcctctgtgataggacacacctgtgctctcctcaaggga	1398
OY	879	accctggatataccatcgggagctgttctgtcgtgttttctttagtgaatccataagcagtta	938
Db	1399	-----atattctatttggaaagatctacactcagctattcttcttgatttgaatttggggcctt	1452
OY	939	ttgcattgtgaagcaacagctccctcacttcttgaacacttcggaatagcccgagggagcgtct	998
Db	1453	tagtgcttggaagcagcattcccaagcatttgaagcatttggaaatgtcaagaaggggcagcctta	1512
OY	999	tcaattttccaaggttatgtataagaacaccagatagataaacttttccacagctgtgata	1058

Db	1513	tgaaatcttcaagaataattgtatataataagccaaglatgtcacagctatctcgaagatgtggca	1572
Qy	1059	taaaccttgaatccataagaagaaactgtggaatttaaaaaagtctcttcaattacac	1118
Db	1573	caaacacgatatataatttaagggaatttggatttcagaaatgttccaattgaattaccac	1632
Qy	1119	aagaccatctataagaattctcgaagagcttgaaatccagaaatgaattcgtcgaagacag	1178
Db	1633	tcgaagaagaagttlaagatctctggaaggccctggaaccctggaaggtgcagagtgtggcagaagct	1692
Qy	1179	cgcccttgctgtctcaatbtgcaggttgagggaagaaatgaagtaagctgaagccagctcttcgaaggt	1238
Db	1693	ggccctgtgttgaaacacaggtgcgtgtgggaaggaacaacagctccagctggaatgcagaagct	1752
Qy	1239	atatatccgagatbatatgtccttatacatgtgtgaatgaatgaacatgaacagagcttlaaagt	1298
Db	1753	ctatgaacccacagagggtgtgtcgaagtgtgtgtgaatgaacagbatatgaacataaatgt	1812
Qy	1299	gcggcatcatcgaagacatatttggaatgtgtltagtcaagaagcctgttttgttcgggaacc	1358
Db	1813	aaggttcttaagggaataatcatctgtgtgtgtgaagtcagaacctgtatttgttgcacacac	1872
Qy	1359	catcaatgaacatbtccaagtatggaagagatgtgtgtgaatgtgaagaagatggaagagc	1418
Db	1873	gatagtctgaaaaacctctgcatactgcgtggaataatgcacacatggaatggaattggaagaagc	1922
Qy	1419	agcaagaggaacaaatgcgtatgatatttatacagtagaggtcttccataaatttaacat	1478
Db	1933	gttcaaggaagccaaatgacctatgtacattatcatagaacacgtccataaaattgacaccc	1992
Qy	1479	ggttaagggaaaaagagactcaaatgtgaatggaagggaaggaacaaagagatcgcgaattgtctg	1538
Db	1993	ggttgaagagagagggcccaagttgaagtgtgtgtggaagcagaagcagaagatcgcacattgcag	2052
Qy	1539	tgcccttgcttgaaaccccaagattctgatttaagtatgagtggttaagtcgtccctgtgattc	1598
Db	2053	tgccctgtgttgcacaccccaagatccctcctgcgtgaatggaagccacgtcagcttggacac	2112
Qy	1599	agaagcgaagtcagctgttccaagctgcagctggaagaag	1636
Db	2113	agaaagcgaagcagtggttccaagtggtcgtcctgcgtgaag	2150
RESULT 15			
AA052726			
ID	AA052726	standard; cDNA; 4669 BP.	
XX	AA052726;		
AC			
XX			
DT	24-JUN-1994	(first entry)	
XX			
DE	Sequence of human multi-drug resistant gene mdrl.		
XX			
OS	P-glycoprotein; multi-drug resistant gene; mdrl; ss.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	424..4267	
FT		/*lag= a	
PN	W09324613-A.		
XX			
PD	09-DEC-1993.		
XX			
PF	14-MAY-1993;	93WO-US0407.	
XX			
PR	22-MAY-1992;	92US-0887712.	
XX			
PA	(GENE-) GENETIC THERAPY INC.		
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
XX	McDonagh KT Nienhuis A, Tolstoshev P;		

```

XX WPI: 1993-405805/50.
DR P-PSDB: AAR44297.
XX
PT DNA or RNA sequence for human multi-drug resistant gene MDRI -
PT encodes P-glycoprotein with altered splice site, used in cancer
PT therapy
XX
PS Example: Fig 4: 64pp: English.
XX
CC PMD2000 contains an mdrl CDNA sequence (AA052726) described in PCR
CC application no. W087/0943m wherein the first 282 bp of the 5' UTR
CC and the last 23 bp of the 3' UTR of the CDNA sequence have been
CC removed. The codon YTN is unclear in the spec., and encodes the AA
CC Leu.
XX
SQ Sequence 4669 BP: 1394 A: 893 C: 1126 G: 1254 T: 2 other:

Query Match      25.3%; Score 511.2; DB 14: Length 4669;
Best Local Similarity 61.9%; Pred. No. 8,6e-142;
Matches 915; Conservative 0; Mismatches 473; Indels 90; Gaps 3;

OY 162 caggttgaccctgtatattgttgaataggctgtgctgtcatttttggtaacataca 221
    ||||| || ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||||
DB 760 caggtatgcctattatatacagtgtaattgtgtcgtgtgtgtgtctgtcattacata 819

OY 222 gatttcctgtgtgattatatacttcagcacagacagacaaaggattccgaacagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 820 gtttcaatttggctgcctgcgcagctcggaagacaatacacaaattagaacacagtttt 879

OY 282 tcattcaagtttggcacagacacacgcgtgtgtgtgtatagctgtgtacatcgcgtgaactaa 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 880 tcctgtatataatgcgcagacagatagagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 939

OY 342 cactcgcatagcagacattgaac---aaatcagtgatgtgattggagataagattgctct 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 940 caccgcgactacagatgatgtctctcaagattaaatgaactattgtgtacaaatttggaat 999

OY 399 ggtgtgtcaaaacatgtcctcttcttcagattggtcctgcgcagctgtgtgtgtgtgtgtgt 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1000 gttccttcacgaatgcagcaacattttcactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1059

OY 459 gaaactaccctatgtgactatccatccacgtctcctctatataatgtgtcgaagcagatg 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1060 gaagctaaccttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1119

OY 519 ttcttagatgtgtatccatcttgacagtaagtaagtaagtaagtaagtaagtaagtaagtaag 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1120 ggcacaaagatctatctcattactatgaataaagaactcttagcgtatgcaaaagctggagc 1179

OY 579 tgtgtgcagaagaagcttctgtcatcaatccgaacatcatagccttggggcccgagagaa 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1180 agtagctgaagagagctctgtgcagcaatagaacatgtgattgtcatttggagacaaagaa 1239

OY 639 agaactcacaaggtcttccctttaaataaacaagaatgtgtgtttatttcccca 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1240 agaacttga-----aaagtaacacaaaatttagaagaagctaa 1248

OY 699 gtgtgctactaagtgtgttctgttntttgttaagtatacagaaatccaaagatgcaaa 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1249 -----aaagtaacacaaaatttagaagaagctaa 1278

OY 759 ggaatttggcataaaagactatagcttcaaaagtctctgtgtgtgtgtgtgtgtgtgtgtgtgt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1279 aagaattgggataaagaagactatacagcaatattctataagtgctgtcttccctgtct 1338

OY 819 taatgaatgaacctatgtgactgtcttctgtgtatggaaacctcctctgtatcttaatgagaa 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1339 gatctatagcatctatgtctgtcgtctctgtgtatggagacacactgtgtcctctcagggga 1398

OY 879 acctgatatccatcggaactgttctgtgtgttcttcttagtgaatccatagacagta 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 1399 -----atatctatggacaagtaactcaactgtatcttcttcgtatttaattggggcttt 1452

OY 939 ttgcatttggacgaagcagcagccctcaactttgaaaccttcgcgaatagcccgagggtgctct 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1453 taagtgtgacagcagcatctccaaagcatttgaagcattttgaaatgcaatgcaagtgagcagctta 1512

OY 999 tcattttccaggttatttgataagaacccaagtatagataactttcccaagcttgata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1513 tgaatcttcaagataatgatataaagaacgaattatgacagctatctgaagaagtgaggca 1572

OY 1059 taacactgaatccatagaagaactgttggaatttaaaatgttcttcaatataccatc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1573 caaacacagataatataatgaaggaaatttggaaattcagaatagtctcacttcagtaaccatc 1632

OY 1119 aagaacatctatacaagatcttcaaaagctcgaatcccaagaataagcttctggaagacagt 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1633 tcgaaaaagaagttaaagatcttgaaggcctggaacctgaaggctgacagatggtggcagacagt 1692

OY 1179 cgcttggctgtcgtcctaagtgcagcttggaagagtagcagtagatccagcttcgcagaggtt 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1693 cgccctgtgttgaaacagctggtcggtgggaagacacacagctccagctgtgacagaggtc 1752

OY 1239 atatgacgcgatgagtgtcttcatatgtgtgaatgagaaatgacatcagagctttaaattgt 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1753 ctatgaccccaagagggagtgatcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1812

OY 1299 gcgcgattatcgaagacatttggagtggttagtcaagagcctgttctgtcgtggaccac 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1813 aaggtttctacgggaaatcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1872

OY 1359 catcagtaacaatatacagatgatgagacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1873 gatagctgaaacacatctgcctatggtgcgtggaatgacacatgatatgatatggaaagc 1932

OY 1419 agcaagggaaagcaaatgcgtatgatttcatcatgagatttccataataatatacatc 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1933 tgtcaaggaagccaatatgcctatgacttattcatatgaactgtccataaatttgcacacct 1992

OY 1479 gttatgggaaaaaaggagctcaaatgtaagtgagggcagaacagagatgcgaattgtctg 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1993 gttctggaagagagggggcccgatgtgagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2052

OY 1539 tgcctttagtctgaagaccccaagatctcgtattttagatgagagctacgtctgtcctggatctc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2053 tgccttgggttgcgaaccccaagatccctcctgtcgtgagtgaagccacgtccagcttggagac 2112

OY 1599 agaaagcagtcagctgttcaagctgcacttggagagaag 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2113 agaaagcgaagcagtggttcaagctgtcgtgtgtatgaag 2150

```

Search completed: April 22, 2002, 23:10:39  
 Job time: 26227 sec



**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:54:17 ; Search time 14771.4 Seconds

(without alignments)  
2257.117 Million cell updates/sec

Title: US-09-873-409-15  
Perfect score: 2021  
Sequence: 1 ttccgcttgctgcatgact.....taaatgcgtactattaga 2021

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hgtgo\_hum:\*  
31: em\_hgtgo\_inv:\*  
32: em\_hgtgo\_rod:\*  
33: em\_hgtgo\_hum:\*  
34: em\_hgtgo\_inv:\*  
35: em\_hgtgo\_rod:\*  
36: em\_hgtgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	533.6	26.4	4279	6	AX105082 Sequence
2	532	26.3	4279	6	AX105078 Sequence
3	532	26.3	4279	6	AX105080 Sequence
4	530.4	26.2	4045	12	AF269224 Synthetic
5	530.4	26.2	4279	6	AX105057 Sequence
6	528.8	26.2	4317	4	AF045016 Canis fam
7	528.8	26.2	4317	6	AX105059 Sequence
8	514.4	25.5	4378	6	E02326 Multidrug r
9	512.8	25.4	4646	6	I49610 Sequence 2
10	512.8	25.4	4646	9	HUMMDR1
11	512.8	25.4	4669	6	AR091275 Sequence
12	512.8	25.4	4669	6	AR091275 Sequence
13	512.8	25.4	6505	6	AR028671 Sequence
14	512.8	25.4	8630	6	AX012321 Sequence
15	512.8	25.4	9318	6	AR028672 Sequence
16	511.2	25.3	8630	6	AX012320 Sequence
17	506.4	25.1	4669	6	AR055785 Sequence
18	504.8	25.0	4186	6	AX108654 Sequence
19	504.8	25.0	4195	6	AX108656 Sequence
20	503.4	24.9	4264	6	AR051647 Sequence
21	501.8	24.8	4264	6	AR051650 Sequence
22	498.4	24.7	3988	6	AX024454 Sequence
23	498.4	24.7	4192	9	AF016535 Homo sapi
24	490.2	24.3	4280	10	CRUPEPIT
25	488.8	24.2	3489	4	AB029153 Felis cat
26	488.6	24.2	4296	10	CRUPEPIT
27	487	24.1	3987	10	CRUPEPIT185
28	487	24.1	4304	10	CRUPEPIT165
29	480.6	23.8	4298	10	MUSMDR
30	477.6	23.6	3858	4	OU78609 Ovis aries
31	477.4	23.6	4233	6	AR123273 Sequence
32	477.4	23.6	4356	10	MUSMDRRA
33	477.4	23.6	4924	10	MUSMDR1A
34	475.8	23.5	4254	10	RATMDRM
35	471	23.3	4018	5	GA09799 Gallus ga
36	469.2	23.2	4323	10	AF266167 Rattus no
37	469.2	23.2	4927	10	AF257746 Rattus no
38	466.4	23.1	4390	5	XI017608 Xenopus lae
39	445.4	22.0	3924	6	AX024455 Sequence
40	445.4	22.0	3924	9	HUMMDR3
41	437.4	21.6	3912	10	RATPELYCO
42	437.4	21.6	4084	10	MUSMDRA
43	432.6	21.4	3905	10	CRUPEPITI
44	412.6	20.4	3682	10	MUSPELYXIA
45	388	19.2	79611	9	AC002486 Human BAC

#### ALIGNMENTS

RESULT 1  
AX105082  
LOCUS AX105082 4279 bp DNA  
DEFINITION Sequence 26 from Patent WO0123540.  
ACCESSION AX105082  
VERSION AX105082.1 GI:13921232  
KEYWORDS  
SOURCE  
ORGANISM  
dog.  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 4279)  
Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Relf,T.C. and  
Patten,C.J.  
P-glycoproteins and uses thereof  
Patent: WO 0123540-A 26 05-APR-2001;  
JOURNAL GENTEST CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..4279  
/organism="Canis familiaris"



ORGANISM	Canis familiaris
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Carnivora: Fissipedia: Canidae: Canis.	
REFERENCE	1 (bases 1 to 4279)
AUTHORS	Stocker, P. J., Stelmel-Crespi, D. T., Crespi, C. L., Reif, T. C. and Patten, C. J.
TITLE	P-glycoproteins and uses thereof
JOURNAL	Patent: WO 0123540-A 22 05-APR-2001;
FEATURES	GENTEST CORPORATION (US)
SOURCE	Location/Qualifiers 1..4279
CDS	/organism="Canis familiaris" /db_xref="taxon:9615" 17..3862 /note="unnamed protein product" /codon_start=1 /protein_id="CAC37721.1" /db_xref="GI:13921228" /translation="MDPEGGRKSGAEENFMKGGKSKKNEKEKKPVVSTFPMRYSNNWLDLVLVGTMAI IHGALPLMLVLFGMNTDSEFANAAGISRNTPVLI INESTTNMT OHGFNLHEEMETTYAAYVYSGIGAGVLA IYQVSEFQCLAGROFLKIRKQFPAIMRO QIENFDVHDVDELNTPLRTFDVSK INEGIGKIGIEAFVGFGLVFTGKRLKLVILIASPVLGSAIIMKILISSFDDKELIARAKGVAEEVLAIRVIAFGGOKVEERKNKNEPAKGIKIKKAITAINISGAALFLITYASVALEEYGLVSLSSPYSGOVLVTFPSYVLGAFSIQASPSIFAPFANAGAAEIRKIIDNRSIDSYSKSGKRPNIKG NLEKKNHFSFPSRKEVLIKGLMKVQSGVTALVQSGGKSTVQLMORLYDPTD GWCIDGQIDLTITNVRHLREITGVVSOEPLYLENTIENIRGENYTMDEIENAVRE ANAVDILKPNKEDTLEVGERGAQLSGGOKRIAIALVLPNKILILDEATSLDTE SEAVVOYALDKARKGRTTIVIAHSLSTRVANDVLAGDDGVLYKGNHDELMKRGLEY FKLTQTRGNEIPLFNATGSGKSESPALMSPDSSSLTKRSTRSIIHAPGODR KLCKPEDLNENIPLPVPSFRILIKLSTEMPRYVQIFCALINGGLOAPASIFESNIICT FTREDEPRKRONSMFSVLEPLVIGTISFTFFLQGFPPGKAGETLTKRLMVFPSRS LROQVSWFDDPKRTTGAITRRLANDAQAQVKAISRLAVITQNLANGTGIIISLIVG WOLNLLLAIPITIAIGAVEMKMSQCALDKRKELGAKIATEALENFTVYSLTRF EOEKEYMAOSILOVYRNSLRKAIIFGSEFISITQAMMYFVACGFRGAYVAFEMPMK FQDLVESAIVFGAMAVGVSSPAPDPAKAKVSAHYIMITEKSPILDYSFPIGLKRM NTELGANTFENVNPNRPDIPIVLOGLSLEVKGQRTLAIVGSGCKSRVYOLISHEE YDPLAKESYLIDGKEIKHLNOMWLAHLGITYSEPILEDCSIAENIATGDNRSVRV SHEE IMQAKENAIHHFTLEPLERKNTVRVGKQTQLSGGOKRIAIALVLRQPHILIDDEE TSADLTSEKVVQDALDKARKRGRTCIYIAHRLSTYIONADLIVVRQNGKVKHGTHOOD LAAQGIYSFMSVQAGAKR"
BASE COUNT	1295 a 833 c 1008 g 1143 t
ORIGIN	
Query Match	26.3%; Score 532; DB 6; Length 4279;
Best Local Similarity	62.8%; Pred. No. 3e-115;
Matches	928; Conservative 0; Mismatches 460; Indels 90; Gaps 3;
0Y	162 caggttgacctgtattatgttggaatagggtgtgtcgtccttgatttttggttacatca 221
Db	358 CACGATGATGCTTTATTTACAGTGGAGATGGTCTCGCGTGGGCGTTCATCATCCA 417
0Y	222 gattcccttggtgattataaactgcagcagcagacgaagaagagatttcgaanaacgtttt 281
Db	418 GGTTCATTCCTTGCTGCTCGCGACGACGAGAACAGATACCAAAATTTAGAAAACAATTTT 477
0Y	282 tcaatcagtlttggtgcacagagacatcgctggtgttgatagctgtgacatcggtgaactaa 341
Db	478 TCATGCTATCATGACGACAGAGATTTGGCTGGTTGACGNCATGACGTTGGGGAGACTTAA 537
0Y	342 cacttcgtcatgaagacatg---acaaaaacgtgatgtgatltggagaataaagtgcct 398
Db	538 CACCCGGGCTCACAGACGAGTGTCTCCAAAATVCATGAAGAAATTTGGCGACAAAATTTGAAT 597
0Y	399 gtgtgttcaaaacatgctcaacttttgcatttgccgtgacccgtgagctgtttgtgaagagct 458
Db	598 GTTCTTTCATCATATGACAACAATTTTTCACGCGTTTAAAGTGGGGTTTACACGTGGTGG 657
0Y	459 gaaactcaccttagtgaactctatccaagctctccctcttataatgcttcagcgagcagcat 518
Db	658 GAAGCTAACCCCTTGATGATTTTGGCCATCAGCCCTGTCTTGGACCTTTCAGCCGCATCTG 717

OY	519	ttctaaagatggtcattctcattgcacgtlaagaaatgaatgctccattctccaaagctggagc	578
Db	718	GGCAAAAGTACTATCTTACTTACTGATGAAGAACTCTTGCCCTATTGCAAAAAGCTGGAGC	777
OY	579	tgtgcagaagaagaatctctgcatcaatccgaacagcatagacttaaggcccgagagaa	638
Db	778	AGTAGCTGAAGAAATCTTTACACGATCAACAACTTGATGTGCTTTGGAGCAAAAGAA	837
OY	639	agaacttcaaaagctcttccttttaaatataacaagatagctgtttatattctccca	698
Db	838	AGAACTTCA-----	846
OY	699	gtgcgtactaaagtgtgctctgtnttgtlaagatatacagaatctcgaagtgcaca	758
Db	847	-----AAGGTACACAAAAAAATTTGAAAGAGCTAA	876
OY	759	ggaatttggcataaagaagcatagcttcaaaagtgctctctgtgctgtactctt	818
Db	877	AGGAATTGGGATTAAGAAAGCTATACGGCAACATTTCATTGCGCCGCTTCTTATT	936
OY	819	tatagatgaacacatctgactgtctcttctgtatcggaacctccttgatctcttaacgaga	878
Db	937	GATCATGCAATCATATCTCTGCGTTTCTGTATGGAGACCTCCTTGTCCTCCAGTGA	996
OY	879	aacttgabataccatcggagactgtctctgtcttctcttaagtgtaaccataagcagta	938
Db	997	-----ATATTCTATTGGACAAAGTACTCACTGCTCTTCTTCTGTATTATTGGGGCTTT	1050
OY	939	tgtcattggagcagcagtcctcctcaacttgaaccttcgaacctcgaataagcccgagagcgcct	998
Db	1051	TAGATTGTGGACAGCATGCCCAAGCATTTGACATTTGGCAACGCAAGAGGAGCACTTA	1110
OY	999	tcaatttccaggttatatgaagaacccaatagataaactttccacagcctgata	1058
Db	1111	TGAAATCTTCAAGTAAATTGTGCATATAACCAACATTTGACAGTATTGCAAGAAGTGACA	1170
OY	1059	taaacctgatalccatagaagaactgtgaaatttaaaaaatgcttcttccaattatccatc	1118
Db	1171	TAAACCAATTAATTATTAAAGGAAATTTGCAATTCAAAAATGTTCACCTTACGTACCCTTC	1230
OY	1119	aagaccatctatcaagttcttgaaagttcgaatccagaattaagtcgggagagaagct	1178
Db	1231	TCGAAAGAAAGTTAAGATCTTTAAAGGGCTCAACCTGAAGTTCACAGTGGAGACAGCT	1290
OY	1179	cgccctgtcggtctcaatctgacagctgggaagatgacgtatgcacgtcttcagagatc	1238
Db	1291	GGCGCTGTTGGGAACAGTGGCTCGGGAAGACGACGCGTCACTGATGTCAGAGCT	1350
OY	1239	atatgatccggaatgatgcttattcatctgctgagatggaatgacatcagaagctttaaagt	1298
Db	1351	CTATGACCCCAACAGATGGCATGCTGTATTGATGGACAGACATTAAGAACCATTAATGT	1410
OY	1299	gcgcgcatatcagagaccatatitgagttgtttagtcaagagcctgttcttctgggagcac	1356
Db	1411	AAGGCATCTTGGGAAATTAATCTGCTGGGAGAGCAGAGACCGCTGTCTTTGGCCACAC	1470
OY	1359	catcagtaacaatcgaatctgagcagatgagatgagatgacgtgaagagatgagagagc	1418
Db	1471	GATACTCAAAAACATTTGCGCTATGGCCCGCAAAATGTCACCATGATGAGATTGAGAAAC	1530
OY	1419	agcaagggaaagcaaatgctgataattatcatitgagtttctctaataattataacatt	1476
Db	1531	TGTTAAGGAAGCAATGCTTGTGATTTTATCATGAAACTACCTAAATAATTGGACACTCT	1590
OY	1479	ggttaagggaaaaagagctcaaatgagatggaagggcagaacaagatatcgcataatctcg	1538
Db	1591	GCTTGGAGAAGAGGGGCCGAGCTGAGATGGTGCAACAATAACAGAGATGCCATTGCTCG	1650
OY	1539	tgccttagtctgaaaacccaagatcttgatatttagaagagctagctctgcgcctgattc	1598
Db	1651	GGCCCTGGTTGCCAACCCCAAGATTTCTTCTGCTGATGAGGCAACGTCACTCTGACAC	1710
OY	1599	agaaagcaagtcagctgttcaagcttgacttggagaag	1636

Db	1711	TGMAAGTCAACGACGGTTCACGTGCGCCCTGATRAAGC	1748
RESULT 3			
AX105080	AX105080	4279 bp	DNA
LOCUS	Sequence 24	from Patent WO0123540.	PAT
DEFINITION	AX105080		30-Apr-2001
ACCESSION	AX105080		
VERSION	AX105080.1	GI:13921230	
KEYWORDS			
SOURCE	dog.		
ORGANISM	Canis familiaris		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
AUTHORS	Stocker,P.J., Steinel-Crespi,D.T., Crespi,C.L., Relf,T.C. and Patten,C.J.		
TITLE	P-glycoproteins and uses thereof		
JOURNAL	Patent: WO 0123540-A 24 05-APR-2001;		
FEATURES	GEMTEST CORPORATION (US)		
Source	Location/Qualifiers		
	1..4279		
	/organism="Canis familiaris"		
	/db_xref="taxon:9615"		
	17..3862		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC37722.1"		
	/db_xref="GI:13921231"		
	/translation="MDPEGRKGSAEKNFMGKSKKREKKEKPPVSTFAMFRYSN		
	WLDRLVVLGTMAIIGGALPLMLVFGMTDFEANAIGSRNTPFVILINESTNTNT		
	OHFNHLEEMETTYAAYVYSGIGAVLVAAYTVQSWFGLAAGROLKIRKQEPHIMRO		
	EIGEDVDHGEMLTRELTDVSKTNEIGIKGIFROSITPPFGITVGRGKRL		
	VIIAISPYVLGLSAIIRAKIISTFDKLLIYAKGAAVEYLARIYVIAFGGKREL		
	ERIKNLNEEAKGIGIKRAITANISIGAFLLITASYVALAFWYGSIVLSSEYSGVVL		
	TFEEVNLIGAFSIGSIOGPSIEAFNANGAAVEIKIIDNKPISIDYSQSHKPDNIG		
	NLEKSNVHFSYPSRKEVKILGLNMLKQSGQVLAIVQNSCGSKSTVOLQRLDPD		
	GNOVDGODIRTIIVRHLREITGVSGQEPVLAFTLLENIRYGRNTEDEIKAVEE		
	ANAVDLMKLPNKFDTLVGRGAQLSGGQORIIAVALRNPKIILLDETSATL		
	SEAVVOYALDKARKGRFTIYIAHRLSTYONADVADGDDVYVGGHDELMKRGIV		
	PKLYVTQGRNEIELENATGESKESDALEKSPDSSLLKRSTRSTHAPOGDPR		
	KLGEDJDNENVPVPSFWRLIKLNTSTWPFVVGIFCALINGGLPAFSIIFSIIGI		
	FTFRODEPTKRQNSMNSVLEVLGIIISFTFPLQGFPGAGSILKRLRYVFRSM		
	LKROVSWFDDPKNTGATLTRLANDAQVAGALGSRVAVTONIANIGTGIISLITG		
	WQULLLAIPIIPIAIGVEMKMLSQALDKLELGAGATIEALENFTVYSLTR		
	EOKEFVYAOGLQVYRNSLRKAIIFVPSFSTIQAMVFSYAGCFRGAYIVAEFNN		
	FODVLVFAIIVFEGMAVGVSSFAFPAKAYAAHVMIIEKSPILDSYSPHGLK		
	NTELEGNVFNENVPNPTPRDIPYQLSLSEVKKQTLALVYSGGCKSTYOLLERF		
	YDPLAGSVLIDGKEIKHLNOMLAHLGIVSOEPLDSCIAENITAVGDSRVVSHFE		
	IMADKEANINHHLETLPKRYNTRVGGKQLSGGQRIATAVALRVQPHIILLDBA		
	TSADTSEKVVQALDKARGRCIVIAHRLSTYONADVADGDDVYVGGHDELMKRGIV		
	LAOKGIYFSMVSVAGAKR"		
BASE COUNT	1296 a	833 c	1008 g 1142 t
ORIGIN			
Query Match	26.3%;	Score 532;	DB 6; Length 4279;
Best Local Similarity	62.8%;	Pred. No. 3e-115;	
Matches 928; Conservative	0;	Mismatches 460;	Indels 90; Gaps 3;
Db	162	caggtgaacctgtatattgttggaatagggtgtgcgtccttgatttttggttaacata	221
Oy	162	caggtgaacctgtatattgttggaatagggtgtgcgtccttgatttttggttaacata	221
Db	358	CACGATATCCCTTAATTTTACAGTGGGATGCGGCTGGCGTGGGCGGTCTTACATCCA	417
Oy	222	gattcccttgatgattataaactgcagcagcagcagcaagaagagattcgaaaaacagttt	281
Db	418	GGTTTCATTTCGTGCTCCCGCGCAGCAGGAGACAGTACTCAAAATTTAAGAAAACAATTTT	477
Oy	282	tcaatcagttttgagcagcagcagcattcgctgtgttttataatagcttgtagcattgaa	341
Db	478	TCATGCTATTCATGACGACAGGAGATGTGGCTGGTTTGACGCGATGACGTTGGGAGACTTAA	537

OY	342	cacccgcatacagagacatctt-----acaaataacagatgatgatcttggagaataagctgacct	398
Db	538	CACCCGGGCTCACAGACAGATGTCCTCCAAATTCATATGAGGAATTTGGCGCAAAATTTGGAAAT	597
OY	399	gtgtcttcaaaaacatgctactcttcttcgaatttgacctgagcttggttttggtgaaggacctg	458
Db	598	gttcttttcaattcaatagcaaacatttttttttcaacgggtttttatagtgggggtttttacacgngttg	657
OY	459	gaaactacccttagtgactctatccacgctctccctcttataatgagcttcaagcgcaagcatg	518
Db	658	GAACCTTAACCCCTGTGTATTTTGGCCATCAAGCCCTGTCTCTGGACATTTTACGGCCGCAATCG	717
OY	519	ttccaggaatggtcatcctcatctgagccgaataaggaatttaagtcctatcccaagcttgggcg	578
Db	718	GGCAAAAGATCTATCTTCTATTTCATATTAAGACTCTTTGGCCTATGCAAAAAGCTGGAGC	777
OY	579	tgttgccagaagaagctctctgtcatcaaatccgaacagtcataagactttaaaggcccgagaa	638
Db	778	AGTAGCTGGAAGAAGTCTTACGACCAATCAGAACGTGATGGCTTTGGAGGCAAAAGAA	837
OY	639	agaactccaaggctcttccctttaaataaacagaatagtcttggttttlatlttcccaca	698
Db	838	AGAAGCTTGA-----	846
OY	699	gtgtgactactaagttggttctgttnttgttaagataacagaaatctccaaggatgcaaa	758
Db	847	-----AAGGTACAACAAAAATTTAGAAAGAGCTAA	876
OY	759	ggaatttggcataaaaagagactatagcttccaaagtgctctgtgtgtgactactctt	818
Db	877	AGGAATTTGGGATTAAGAAAGACTTCAAGCGCCAAATTTCTATTTGGTGGCGCTTCTTATTT	936
OY	819	tatgatggaacctatggaacttgcttttcttgatataggaacctcttgacttcttaatgagaa	878
Db	937	GATCTATGACATCATATGCTCTGGCTTTGATGATGAGGACCTCTTGGHCTCTCCAGTGA	996
OY	879	acctgagataaccatcggaactgttctctgtcttcttcttagtgaatccatagacagtta	938
Db	997	-----ATATTTCTATTGGAACAAAGTACTCACTGCTCTCTTTCTGTATTAATTTGGGGCTTT	1050
OY	939	tgtcatgtggagcaagctccctccactcttgaaccttgcgaactgcgaatagccagagagctgacct	998
Db	1051	TAGATTTTGGACAGGCAATCCCCAAGCATTAAGCATTTTGGAAACGCAAGAGGACAGCTTA	1110
OY	999	tcatatttccaggttatatgataagaaccagctatagataacttccacagctgagata	1058
Db	1111	TGAATCTTCAAGATATATTTAGCAATTAACCAAGCATTTGACACTTATTTGGAAGAGTGACA	1170
OY	1059	taaacctggaatccatagagaagaactgttgaaatttaaaaatgttcttccaattatccatc	1118
Db	1171	TAAACCGAATATATTTAAGGGAAATTTGGAAATTCAAAATATGTTCACTTCAAGTTACCTTC	1230
OY	1119	aagaccatctatacaagaattctgaagaagcttgaatcaccgaatttaagcttggagagacagt	1178
Db	1231	TCGAAAAGAGATTAAAGATCTTTAAAGGCTGCAAACCTGAAGTTCCAGAGTGGGGCAGACAGT	1290
OY	1179	cgccttggctgcgttctcaatggcgatggaagagtagcgttagtccagcttctgcagaagtt	1238
Db	1291	GGCGCTGATTGGGAACAGTGGCTGCGGGAAAGACAGACCTGTGACGTGATGACAGAGGCT	1350
OY	1239	ataatgaccggatgtagtgctttatcatgtytgatgtagaattgacatccagagctttaaagt	1298
Db	1351	CTATAGCCCCCAATATGGATGGTGTCTGTATTTGATGAGCAGGAGCATTTAGGACCATTAATAGT	1410
OY	1299	ggcgcatatccagagacaatttggagtggttgaataagaagcctgttcttgctcggagcaac	1358
Db	1411	AAGGCATCTTTCGGGAATTTCTGCTGTGTGTGACTAGGAGCCTGTGTGTTTGGCACCAC	1470
OY	1359	catcagataacaatatacgaatataggacagagatgtagtgcagatgtagaagatggaagagc	1418
Db	1471	GATAGCTGGAAGAACTTTCGCTATAGGCGCGGCAAAATGTACACATGAGATGAGATGAGAAAGC	1530
OY	1419	agcaaggaagcaaatgycgatatgataatttatcatcagtagtcttccataaattatacaatt	1478

DB	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	AUTHORS	TITLE	JOURNAL	FEATURES	BASE COUNT
Db	1531	TTTAAAGACCAATGCTCTATGATTTTATCATGTGAACCTACTTAATTAATTGGACACCT													1171 a 850 c 986 g 1038 t
0y	1479	ggttaggggaaaaagagatcgaatagtgagggagggagaacagagatcgcgaattgctcg													
Db	1591	ggttgagagagagaggggccacgactgagtgatgagagacagaaacagaaatccgcatggctcg													
0y	1539	tgccttagtgcgaaccccaagatctgatttagtagaggtcagctctgcctgattc													
Db	1651	ggcccttggtgcgaaccccaagatctctctgctgagtagagcaacgacgctgcgacac													
0y	1599	agaagcaagtcagctgttcgaagcgcgcacgtggaagag													
Db	1711	tgaaagtgacagcagtggttcaggtgcccctgcatgagc													
RESULT	4														
AF269224															
LOCUS	AF269224	4045 bp	DNA	SYN	04-JUL-2000										
DEFINITION		Synthetic construct Canis familiaris his-tagged-multidrug													
ACCESSION		resistance glycoprotein gene, complete cds.													
KEYWORDS		AF269224													
ORGANISM		AF269224.1	GI:8926216												
REFERENCE															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															
JOURNAL															
AUTHORS															
TITLE															

Query Match	26.2%	Score 530.4	DB 12	Length 4045
Best Local Similarity	62.7%	Prod. No. 7.2e-115		
Matches 927	Conservative 0	Mismatches 461	Indels 90	Gaps 3
QY 162	caggttcagccctgattatgtttgatagatggtctgtcgcctgattcttggttacataca	221		
DB 460	CACGATATGCGCTATATTATACAGTGGGATGGGTCTGGCTGGCTGGTGGCTGCTTACATCA	519		
QY 222	gattctcctgtgattatataactctgcagcagcagacccaagagattcgaacagcttct	281		
DB 520	GCTTTCAATTCTGGTGGCTCCGCGACGAGGAAGACAGATACCTCAAAATTAAGAAAAATTTT	579		
QY 282	tcatcagttttggagcagggacatcgcctgtgtttatagcvtgtaacatcgtgaaactaa	341		
DB 580	TCATCTCATCATCGACACGGAGATGGCTGGTTTACGTGCATGCTGTTGGGACCTTAA	639		
QY 342	cactgcatacagacacatctg---acaaaatcagltagtgtlctggaagataagattgctct	398		
DB 640	CACCCGGGCTACAGACGATGTCTCCAAATATCAATGAAGAAATTGGCGACAATAATGGAAT	699		
QY 399	gttgtttcaaaaacatgtctactcttcttcgatctggcctggcagttgttgtgtgaaggtctg	458		
DB 700	GTTCTTTCAATCAATAGCAACAATTTTACCCGGTTTTATAGGGGGTTTACACGTGGTGG	759		
QY 459	gaaactaccctcagtgagactctatccacgtctctctcttataatgcttgaagcagcagctg	518		
DB 760	GAGGTAAACCTCTTGATTTTGGCCATACGCCCTGTCTTGACACTTTACGCCCATCTG	819		
QY 519	ttctagatgtgtcatcctcaatgtaccagtaagaattaaatgacctatccaaagctggggc	578		
DB 820	GGCAAGATACATATCTTATTTACGATAAAGAACTCTGGCCCTATGCAAAAAGCTGGAGC	879		
QY 579	tgtgcagaagaagctctgtcatcatccagacagctcagcttaaggccaggaagaa	638		
DB 880	ATACTCTGAAGAGTCTTTAGCAGCATCAAGAACTGTGATGCTTTGGAGGACAAAGAA	939		
QY 639	agaacttcaaaagctcttcccttctaatacaaaagatatgttgtttatatttcccaca	698		
DB 940	ACAACCTTGA-----	948		
QY 699	gtgcctactaagtgtgtctgtcttcttgttaagtatcacagaatctcacaagtgcacaa	758		
DB 949	-----AAGGTACAAACAAAATTTTAAAGAAAGCTTAA	978		
QY 759	ggaatttggcacaanaaaggacataagcttcaaaagtgctctgtgtgctgttactctt	818		
DB 979	AAGAATTTGGGATTAAGAAAGACGTATACGGCCAACTTTCTATTGGTGGCCGCTTTTATT	1038		
QY 819	tatgaattggaacctatggaactgtctcttctgtatgtaagaaacctcttgattcttaatgaga	878		
DB 1039	GATCATGTCATATTGCTCTGGCTTTCGGATGGGACCTCTTGCTCTCCAGTGA	1098		
QY 879	accttgatataaccatcgcggagcgtctctgtctgttcttctttagtgaatcaatagaagta	938		
DB 1099	-----ATAATTATTATGGACAAGTACTCACTGCTGTCTTTTCTTGATTAATTTGGGGCTTT	1152		
QY 939	tctgcatctggagcagcagctccctcactcttgaacaccttgcataagccagagagcgcctt	998		
DB 1153	TGATTTTGGACACGGCATCCCCCAAGATTTGAACATTTTGCMAACGCAAGAGGACGTTA	1212		
QY 999	tcaatatttccaggttatattgatalaagaaacccagtatagataacttcccaagcttgata	1058		
DB 1213	TGAAATCTTCAAGATTAATTGACAAATTAACCAAGCATTTGACAGCTTTTGGAGAGGTGACA	1272		
QY 1059	taaactgaatcatcagaagaagcgtgtgaaattttaaagtcttcttcaatcatcactc	1118		
DB 1273	TAAACCAAGTAAATATTAAAGGAAATTTGGCAATTTCAAAAATGTTCACTTCAGTTACCTTTC	1332		
QY 1119	aagaccatctacaagattctgaaagtgctgtaatctcagaattaaagctgtgaaagacgt	1178		
DB 1333	TGGAAGAAGATTAAAGGCTTAAAGGCTTCAACCTGAAGGTTTCAAGAGTGGGCAACAGT	1392		

OY	1179	cgccttgatgctcgaattcaatggacgttbgagaagttacgtaacccagctcttcgaagatt	1238
Db	1393	GCGCCTGTTGGGAACAGTGGCTGCCGGAAAGCACAGACCCTGCAGCTGATGACAGGCT	1452
OY	1239	atacgatccgatgatgctgactttaacaagglygatgtgaagtgaatcatccaagctttaaatgt	1298
Db	1453	CTATGACCCCAACAGATGGCATGTGTCGTATTGATGACAGSAGATTAGAACATTAATGT	1512
OY	1299	gcgcgatalatcgagaccatatcttggagttgattgattcaagaagcctgtttgttcocggagcac	1358
Db	1513	AAGGCATCTTCCGGGAATTAACAAGTGCTGGTAGTCAGGAGCCTGGTGTGTTGGCCACAC	1572
OY	1359	catcaatgaacaatatcaagatctgagcagatgtagtgcata tgaagaatbtgaagagc	1418
Db	1573	GATAGCTGGAAAACACTTCCGTCATGTGGCCGCCGAATAATGCACACATGAGTAGTGGAAAGC	1632
OY	1419	agcaaggaaagcaaatgcyglatgatttatcatgatgaagtttctaataaattaalacatt	1478
Db	1633	TGTTAAGGAAGCCAMTCCCTCATTTTATCATGAAACACTCCATTAATAATTGACACTCT	1692
OY	1479	ggttaggggaaaaaaggagcttaaaatggatggagggcagaagaagaagatcgcgaattgcgcg	1538
Db	1693	GCTTGGAGAAGAGGGGCCACGCTGAGTGTGACACAGAAACAGAGAAATGCCATTGCTCG	1732
OY	1539	tgccttagttcogaaaccccccaagaatttcygaattttagatggcctagctcgcacctgattc	1598
Db	1753	GGCCCTGTTGCGAACCACCCCAAGATTTCTTCGCTGATGAGCGCAACGTGAGCTCTGCACAC	1812
OY	1599	agaaagaacatgcagctggtttaaactctgcactcgtgagaag	1636
Db	1813	TGAAGTGAACAGTGGTGTATAGTGGCCCTCGGATTAAG	1850

RESULT	5			
LOCUS	AX105057			
DEFINITION	AX105057 Sequence 1 from Patent WO0123540.			
ACCESSION	AX105057			
VERSION	AX105057.1			
KEYWORDS	GI:13921209			
SOURCE	dog.			
ORGANISM	Canis familiaris			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
AUTHORS	Stockert,P.J., Stelmel-Creepi,D.T., Creepi,C.L., Reif,T.C. and Patten,C.J.			
TITLE	P-glycoproteins and uses thereof			
JOURNAL	Patent: WO 0123540-A 1 05-Apr-2001;			
FEATURES	GENESTEST CORPORATION (US)			
SOURCE	Location/Qualifiers			
	1..4279			
	/organism="Canis familiaris"			
	/db_xref="taxon:9615"			
	17..3862			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC37718.1"			
	/db_xref="GI:13921210"			
	/translation="MDPEGGRGSAENFNMCKKSKNKKRPPVSTTFAMFRYSNNWLDLVMVLTGTMALIHGALPLMLLVGNWTDSPFANAGISRNTFPVJINESTNTNTMFOHFNHLEEMTYYAAYYGAGVLAAVYQVFPFSLAGRQLTKRKQFPAHMFQEIQGFVDHVGELNTRLDLDPSKINEGICQIGFPHSIATFPFGLVGFGRKLLTVLLIASPVLGSAIMIAKTKITSSPTDKLLAYAKGAAVEELAIRVYAFAGCKRELERYKNLNEBAKGITIKKAITDNISITGAEFLITASVLAENYGTSLVSSSTIGQVLTFEFSVLIGAFSISQASPSLEAFANAGAAVEIKIIDNKRPSIDSYSGKHGPNINIGTNEERKNVFSYPSRREVKILGMLKYOSGQTVLVNNSGCKSTVOLQMRLPDNDGMVQIDGQDITINVRHLREITGVSOEPLYATTIENIRYGRENTMDEIKAVEKEKSEANVQALDKARKGRRTTYIAHSLSTVRNADYAGDDGYIEKGNHDEMKRGIYIFKLTVQNGEIELENAVNTGSKESQASPSLSLTKRSTRSIHAPQGDORKLTKRQEDLNENVPVSWRLIKNTSMPFVVGIVCAIINGGLQPFASISFSTIGDI			

BASE COUNT 1294 a 834 c 1008 g 1143 t

ORIGIN

	Query Match	26.2%	Score 530.4	DB 6	Length 4279
	Best Local Similarity	62.7%	Pred. No. 7.3e-115		
	Matches 927	Conservative 0	Mismatches 461	Indels 90	Gaps
QY	162	caagttgaccctgataatgattgtagaagatggtctgccttgattttggtataca	221		
DB	358	CACGTATGCCATTATATACAGTGGATTCGGTCCGTGGCGGTGGTCTTACATCCA	417		
QY	222	gattcctctgtgataataactctgacacagacagaccagagatctgaaaaagttttc	281		
DB	418	GGTTTCATTCGTGGGCTGGCAGCACAGAAAGACAGACTCAAAATTAACAAACATTTT	477		
QY	282	tcattcagttttgacagagacacgcgtgtgtttgataagctgtgacacggtgaaactaa	341		
DB	478	TCATGCTATCATGCGACAGAGATGTGGTGGTTGACGTGATGACGTTGGGGAGCTTAA	537		
QY	342	cactgcgaatgacagacatg----aacaatcagtgatgtgatgtgaagataatgtctc	398		
DB	538	CACCCGGCTCACACACCATGTCTCCAAATTCATAGAAAGAAATTTGGCGCAAAATTTGAAT	597		
QY	399	gtgtttccaacaacgtgctactcttttcagattgcccgtgagttgtgtgtgaaaggtcg	458		
DB	598	GTTCTTTCACGCATATACCAACATTTTTCACCGTTTTATAGTGGGGTTTACACGTGGTTG	657		
QY	459	gaaacacaccatgtagctctatccacgctccctcttaaatgcttcaagcggacgat	518		
DB	658	GAACGTAAACCTTGTGATTTTGGCCATCAGACCCCTGTTCTTGACCTTTCAGCGGCATCTG	717		
QY	519	ttctagaatgtcatctccatctgacagtaagaattaaagtgctatcttccaaagcttgagc	578		
DB	718	GGCAAGATTACTATCTTCATTTACTGATTAAGAACTCTTGGCCTTATGCAAAAGCTGGAGC	777		
QY	579	tgtgcaagaagaagctctctgtgacacaaatccgaacagatagactcttaagggccagagaa	638		
DB	778	AGTAGCTGAAGAGCTTTAGCACACATTCAGAACTGATGTGCTTTGGAGGACAAAGAA	837		
QY	639	agaactccaagaatcttccctttaataataacaagatatgctgtgtttatttcccca	698		
DB	838	AGAACTTGA-----	846		
QY	699	gtgctactaagtgtgttctgttnttgytaaggtatcacacgaatctccaagaatgacaa	758		
DB	847	-----AAGTACACAACAAAATTTTGAAGAAAGCTAA	876		
QY	759	ggatttggcataaaagaagctatagcttccaagaatgctctgtgtgctgtgtactctc	818		
DB	877	AGGATTTGGGATTAAGAAAGAACTATCACAGGCAACATTTTATTTGGTGGCGCTTCTTATTT	936		
QY	819	tatgaatggaacctatgactgtcttttggatagtgaaacctctgtatcttaatgtagaa	878		
DB	937	GATCTATGCAATATATCTTGCTGGCTTTCTGATATGGAGACTCTTGTCCTCTCCAGTGA	996		
QY	879	aactgtgataacacatcgagactgttctgtcgtgtttctttagtgttaatccatagcagtta	938		
DB	997	-----ATATCTATTTGGACAAGACTCACTGTCCTCTTCTGTATTAATTTGGGGCTTT	1050		
QY	939	tgtcaatggagcagcagctccctcaactttgaaaccttcgcaatagcccgagagctgtcctt	998		
DB	1051	TAGATATGGACAGGCATTCGCCAAGCATTAAGCATTTTGAAGACGCAAGAGGACGCTTA	1110		

OY	999	tcataattccaaagttatgtataagaaccccaatgatataactttccaaacttgata	1058
Db	1111	TGAATCTTCAAGATAATTAATGACATTAACAACGAATTCACAGCTATTCGAAGAGTGCACA	1170
OY	1059	taaaccttcaatccatagaagaactcgttgaatttaaaaatgtcttcttcaattatccatc	1118
Db	1171	TAAACCAATATATATTAAAGGGAATTTGGAGATTCAAAAATTTCCACTTCACATTACCCCTTC	1230
OY	1119	aagaccatctatcaagaatcttgaagaagtcgtaacctccagaattgaattcttgaagaacgt	1178
Db	1231	TGCAAAACAATTTAAGACTTTAAAGGGGCTCCACCTGAAGGTTTCAGATGGGCGACAGAGT	1290
OY	1179	cgacctgttcgtctccaaatggcgaatgggaagaagtaagtaacagcttctgcagaagtt	1238
Db	1291	GGCCGTGTTTGGGAACAATGGCTGCGGGAAAGCAGCCGTCGACGCTGATGACAGAGGCT	1350
OY	1239	atattgatccgaatgatagcttatacatctgttgaatgagaatcacatcagaagctttaatgt	1298
Db	1351	CTATACACCCACAGATGGCATGCTGCTGTATTTAGTGAGACAGACATTAAGGACATTAATATGT	1410
OY	1299	gcgcgcatatctgagagaccatattggaagtgttaagtaagaagccgttttgtctcggaccac	1358
Db	1411	AAGGCATTTCTGGGAAATTACTGCTGCTGCTGATGACGAGAGCCTTGTTGTTGGCCACACAC	1470
OY	1359	catcagtaacatatcaaatatgtgaacgaagatcgtatcgtatgaagaatgagagagagac	1418
Db	1471	GATAGCTAAAAACATTGCCCTATGGCCCGGAANAATGTCACCATGGATGAGATTGGAANAAGC	1530
OY	1419	agcaagggaagcaaatgcgtatgtattatcatcgtgaagtttccataaattatlaacatt	1478
Db	1531	TGTTAAAGGAACCCCAATGCCCTATGATTTTATCATGAACATACCTATTAATTAATTGGACACTCT	1590
OY	1479	gttaagggaagaaagaagcttcaaatggaagtggagggcagaagaagaaggttcgcaattgtctcg	1538
Db	1591	GCTTGGAGAGAGAGGGGCCCTACGCTGAAGTGGTGGACAGAAACAGAGAAATCGCCATTGGCTCG	1650
OY	1539	tgaccttaattcgaaaaccccaagaattcttgattttagatgaagagctcgtcgccttgattc	1598
Db	1651	GGCCCTGTGTCGAACCCCAAGATTTCTTCTGCTGGATGAGGCACAAAGTCAGCTCGAGACAC	1710
OY	1599	agaaagcaagtcagactgttcaagcttgcaactcgtgagaagaag	1636
Db	1711	TGAAGTGAACGATGCTTTCAGGTGGCCCTCGATGAAGAAG	1748

[illegible]

```

BASE COUNT      1293 a      844 c      1019 g      1101 t
ORIGIN
/codon_start=1
/product="multidrug resistance p-glycoprotein"
/protein_id="AAC02113.1"
/db_xref="gi:2852441"
/translation="MDPEGGRKSGAEKNFMKGRKKSEKKKKKPTVSTFAMFYSNMW
LDRLXLMVGTMAAIIHGCAALPLMLVLVGNMTDSRANNGISNKKFPVILINESTINRQ
HFNIHLEEMTTVAIVYSGIAGVLAAYIOWSTWCILAAGQIIILIKRQFHALMRQD
IGMVDHVEGVLNTRLTDNDVSGINEGIGDGMFSGTAFEGPITAFEGPMLVLRL
ILAIISPLVGSAAIMAKIILISFTDQKLELLAAKAAVAEEVLAALRTVIAFEQGGKLELE
RYNKLLEAKRIGITKKAITANISTGAAILVAASVLAIVSTVIASEETIGCVLT
VFESVVLIGAFISGASPSIEAFANARCAITEIFITIDNKPISDYSKSGHKPDIAKGN
LEFNHVSYPEKREVKILKGLNKVSGQVVALVNGSGCKSTVYDMLRDXDPTGGS
MVCIDGODIRIINVRHLRELTGVVSGPEVLFEATFIANIRGRENMTDETEKAEKAE
NAYDFIKMLPKRPDLTIVGECARGLSGGOKOIALARLVNRPKILLDEATSDPSES
EAYOVAULDKARKRRTTIVIAHRLSTYRANADVIAGFDGQVIVEGNHDELKKEGITYF
EALVMTQRTGNEIELENAIGESSESDALEMSPKSGSLIKRRTSRNIIAHPQGDOKR
LGTEDEDLENENPVSFMRILKINSTEMPEVWVGJFCALINGLOPASIISFRIGIPE
TRDDDEPTEKRONSNMPSVPLFLVLTISITPEFLOGFSGKGETILRKLRYMVRSMIL
KODYSWEDDPKNTTGALTLLRLANDAAYOKGATGRSLAVITTONIANLSTGIISLYEH
QUTLLLAIVPIAIAGVENKMKILSGALKKORKELEGAKTATATIAENFRVSLTNEE
OKFPEYMAOSLOVPIRNSLRKHLIFGVSTTQMAWTFSGCGCFPACVIAENFVSLTNEE
ODVLTLSAIVEGMAAGVOSFADPDIKAVASAHAYIMILIERPGLIDVSPHLKRN
TLEBNVTNFVEYVNPYRPOIPIVQIGLSLEKVKKOTLATLVNSSCGSLSVOLLERPY
DPLGASVILIDSEKIKHLNVQMLRAHLGLVSEPELLPDCISIAENIAYCDNSRVSHEED
MOAKENANIHHFIETLPKQNTIRVGDGTQISGGOKORIALARLVNRQPHILLDEQNT
SALDTSEKIEVQOALDKARBGRTCTIVIAHRLSTIONADLVVFPNGKVKEHGTQOALL
AOKIYPSMISVOGAKR"

```

Query Match	26.2%	Score 528.8	DB 4	Length 4317
Best Local Similarity	62.7%	Pred. No. 1.7e+114		
Matches 926	Conservative 0	Mismatches 462	Indels 90	Gaps 3
Qy	162	caggttgcacctgtatattgltgaaataagtgltgctgcctgtatcttctgtgtaataca	221	
Db	408	CACGATGCGCTATATATACAGTGGGATGGTCTGGCGTGGTGCTTACATTC	467	
Qy	222	gattcccttgtagtataactgcgcgcacgcagacccaagagatctgaaaaagttctt	281	
Db	468	GCTTTCATTTCTGTGTCCTGGCAGCAGGAAGACAGATACTCAAAATTTAGAAAATTTT	527	
Qy	282	tcattcagtttgcgcacagacatgcgcgtgttataagcttgacatcgcgtgaaactaa	341	
Db	528	TCATGCTATCAAGCAGACAGAGATATGGCTGGTTACGTGCATGACGTTGGGAGACTTAA	587	
Qy	342	caatcgcatagcacagacatg---acaaaatcagtgatgltatgtagaataagttgctc	398	
Db	588	CACCGGCGCTACAGACGATGTCTCCAAATCATATAGAAATTTGGCGACAAATGTGGAT	647	
Qy	399	gttgtttcaaaacatgcttaacttcttcgtatgctgcgcagttgtgttgtagaagagtg	458	
Db	648	GTTCTTTCAATATAGCAACAATTTTTCACCGGTTTATAGTGGGGTTTACACCTGGTTG	707	
Qy	459	gaactcacccatgtagactctatccacgctctccctctlaatgcttcaagcgacagtg	518	
Db	708	GAGCGTAAACCTTGTGATTTTGGCCATGACGCCCTGTCTGTGACATTTACAGCCCATGTG	767	
Qy	519	tcttgtagtgcgtcatcatatagcacaagcaagaatlaagtgctctatccaagctggggc	578	
Db	768	GGCAAGATACTATCTTCAATTACTTATGATTAAGAAACTCTTGGCTTATGCAAAAGCTGGAGC	827	
Qy	579	tgtagcagaagaagctctgtgtcalcaatccogaacagtcataagccttlaagcccaagaa	638	
Db	828	AGTACTCGAAGAAGCTTTAGCAGCAATCAGAACTGTGATTTGCCCTTTGAGAGCAAAAGAA	887	
Qy	639	agaacttcaaaagtcttctcctttaataatacaagaatagtggtttattttcccca	698	
Db	888	AAAACTTGA-----	896	
Qy	699	gtgcgtactaaagtgtgtctgtlntttgtaaggtatacacgaactctcaaaagtcaaa	758	



```

Db 897 -----AAGTACAAACAAAAATTAGAGACGCTAA 926
QY 759 ggaattggcataaaaggactatagcttcaaaagtcgtctgtgctgtgactctt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 927 AAGATTGGATTAAGAAAGCTATCACGGCAACATTTTCATTGGGCCCTTTCTTAT 966
QY 819 tatgaatggaacatgagactgtcttctgtatggaacccctctgacttcaatgaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 987 GATCTATGATCATATGCTGCTTCTGCTGATGGAGACCTCTGCTGCTCTCCAGTGA 1046
QY 879 acctgatatatacatcggaactgtctctgtcttcttcttaagtgatatacatagca 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1047 -----ATATATATTGGACAGGTAAGTCTGCTCTTTCTTTCTTTCTTTCTTT 1100
QY 939 ttgcatgtgagacagcagtcctcacttctgaaaccttcgaatagcccgagagctgctt 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1101 TAGTATTGGACAGGATCCCAAGCATTTGAGATTTTGCAAGCAAGAGAGACAGCTTA 1160
QY 999 tcatatttccaggttatatgataagaaccaglatagataacttttccacagctgata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1161 TGAATCTTCAAGATATATGACATTAACCAAGCATTTGACAGCTATTCGAAGAGTGACA 1220
QY 1059 taacctgtaacatagaagaactgtgaaatttaaaatgttcttcaatatac 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1221 TAAACAGATTAATTAAGGAAATTTGAAATTCAAAAATGTTCACTTCACTTACCTTC 1280
QY 1119 aagacacatcataagaacttctgaagctcgaatctcagaatgaactctgagagacat 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1281 TCGAAAAAGAGTTAAGTCTTTAAAGGCTTCMACTTCAAGCTTCAAGCTTGGCGAGACAGT 1340
QY 1179 cgcctgtgctgtcctcaatgagcagtgagagaglacgtagtccagctctctgcagaggtt 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1341 GGGCGCTGTTGGGAACAGTGGCTGCGGGAAGACAGCAGCCGTGATGACAGAGCT 1400
QY 1239 atctgacccgagatgagcttcaatgagtgagtgagaaatgaatgaatgagcttaaat 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1401 CTATGACCCCAAGATGCAATGGCTGTATGATGAGACAGAGATTAAGACATTAATGT 1460
QY 1299 ggcgcattatcagaacacatctgagtgagtgagtgagtgagtgagtgagtgagtgag 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1461 AAGGATCTTTCGGGAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1520
QY 1359 catcagtaacaatatcaagatagagagagatgagtgagtgagtgagtgagtgagtgag 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1521 GATAGCTGAAACAACTTCGCTATGCGCCGGAATGTCCACATGATGAGATTGGAAGAC 1580
QY 1419 agcaagaaggaacaaatgctgatagttattatcatgagtgagtgagtgagtgagtgag 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1581 TGTTAAGGAAGCCAAATGCTTATGATTTTATCATATAAATTAATTTGACACTCT 1640
QY 1479 gtagggggaaggaagagctcaaatgagtgagtgagtgagtgagtgagtgagtgagtgag 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1641 GGTTCGAGAGAGAGAGGCGCGCTGAGTGTGAGACAGAAACAGAGAAATCGCATTTGCTCG 1700
QY 1539 tgccttagtgcgaacccccaagatctgattttagatgagtgagtgagtgagtgagtgag 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1701 GGCCCTGCTTCGCAACCCCAAGATTCTTCCTGATGAGGCAACGCTCAGCTGAGACAC 1760
QY 1599 agaaagcaagtcagctgttcaagctgagtgagtgagtgagtgagtgagtgagtgag 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1761 TGAAGTGAAGCACTGTGTGAGTGCCCTGATTAAG 1798

```

```

RESULT 7
LOCUS AX105059 4317 bp DNA PAT 30-APR-2001
DEFINITION Sequence 3 from Patent WO0123540.
ACCESSION AX105059
VERSION AX105059.1 GI:13921211
KEYWORDS
SOURCE dog.
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 4317)
AUTHORS Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
Patten,C.J.
TITLE P-glycoproteins and uses thereof
JOURNAL Patent: WO 0123540-A 3 05-APR-2001;
GENEST CORPORATION (US)
FEATURES
source Location/Qualifiers
1..4317
/organism="Canis familiaris"
/db_xref="taxon:9615"
70..3912
/feature="taxon:9615"
CDS
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAC37719.1"
/db_xref="GI:13921212"
/translation="MDPEGKRGKSAEKNFMKMGKSKKKEKKKPTVSTFAMFRYSNM
IDRLVWLVTCTMAITIHGALPLMLVFGMSTDSFPAAGISRNKTFEVLINESTINNO
HFNNHLEEMTYIAYYSIGAGVLYAATIQVSEFCLAGROILIKRQFPFAMRQE
IGFEDVDVDELNTRLDVSKINBGIKGVGFROSITATFTFTGTYGTPRQKILY
ILASPVGLISAIIMAKILISFDKELAVARAGAAVEVLAIRTVAFGQKLE
RYNKLLEAKRIGIKKAITANISIGAFLLIYASVALAFWYLSVLSSEYITGOVLT
VFESVLIGAFSIGOASPSIEAFANAGAAVEIFKIDNRPISIDYSKSHKPDNIGN
LEKRVHPSPSRKEVKKILGLNLKVOSQYVALVNGSGSTVQLMORLYDPDNG
MCTIDQDITRTIYRHLRETGVYSEPLFATTAENIRGRENVTMEIKAVEA
NAVDFIMKLPMKFDTLVBERGALSGGQRIATARALVNRKILLIDATSLDYES
EAVVOVALDKARKGRPTIYIAHRLSTVRNADVAJGDVIVKGGHDLMEKGIYE
KLVTMOTRNEIELENATGESKESDALMSPDSLSLKRSTRSJHAPGDQDK
LGTKEDLNENVEFVSFMRILKLNSTGEMFVFCALINGLOPAFSSIFSRITQIF
TRDEPBTNRONMFMFVLYVIGIISPTFELGTFPGAGEILKRLRYWFRML
RODVSFDDPKNTGATLTRLANDVADVAGISRLVATONANTGSLIISLYGM
OLLLALIAPIITAIAGVBYEMKLSQALDKKDELDACKIAEALENRIVYSIYRE
OKEEYVYASLOVYRNSLRKAHIFGVFSITDAMVYISAGCFRGAIVANEPMN
TEGANTEFNEVYENPTRPDIPLYGLSLEVKKGQYLAIVSGSGCSKYVYDLERY
DPLAGSVLDIGKEIKHLNOMWLAHLGIYSOGLIFDCSIAENIYAGDSMYSHEBI
MOKAENAHNHFTEILPEKYNTRVGDKGLOSGGQRIATARALVROPHILIDAT
SADTSEKYOVALDKARKRGKICYIAHRLSTIQNADLIVFQNKVHEHGHQDL
AOKGTFMSIYVQAGKR"
BASE COUNT 1293 a 844 c 1019 g 1161 t
ORIGIN
Query Match 26.2%; Score 528.8; DB 6; Length 4317;
Best Local Similarity 62.7%; Pred. No. 1.7e-114;
Matches 926; Conservative 0; Mismatches 462; Indels 90; Gaps 3;
QY 162 caggttgacccctgattatgttgaatagtggtgctgagcttcttctgatacatata 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 408 CACGTATGCTATTATTAACAGTGGATCGGTGCGCTGCTGCTGCTTCAATCCA 467
QY 222 gatttccttgataltatcagtcagcagcagcagcagcagcagcagcagcagcagcagc 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 468 GGTTCATCTTCGTGGCCCGCAGCAGAGACAGATTAATCAAAATTAAGAAAATTTT 527
QY 282 tcaatcagtttgacacagacacacagcagcagcagcagcagcagcagcagcagcagc 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 528 TCATGCTATCATGCGACAGAGAGATGCTGCTGCTGACATGATCACTTGGGGAGCTTAA 587
QY 342 cactcgcatacagacacatcttcaatcagtgatgagtgagtgagtgagtgagtgagtgag 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 588 CACCGGCTCAAGACGATGCTCCAAATATCAATGAAGAAATTTGGCAGCAAGATTTGAAT 647
QY 399 gttgttcaaacatgtctacttcttctgattgctcagcagcagcagcagcagcagcagcag 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 648 GTTCCTTCAATCAATAGACATTTTTCACGGGTTTATGAGGGTTTACACCTGGTTG 707
QY 459 gaaactcacctagtgactatccacagctctcctctataatgagtgagtgagtgagtgag 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 708 GAAGCTAACCTTGATTTGGCCATGACCCCTGCTTCTTGACTTTCAGACTTTCAGCCCATCTG 767
QY 519 ttctagatggtatctatctatgaccagtaagaatgagtgagtgagtgagtgagtgagtgag 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 768 GGCAGAGTACTACTTCTTACTGATTAAGAACTTGGCCATGCAAAAGCTGAGC 827

Qy 579 tctgcaagaagaagctctgtcatcatccgaacagctcagctttagggccagagaa 638

Db 828 AGTAGCTGAAGAAGCTTAGAGCAATCAGAACTGTGATGGCTTTGGAGGCAAAAGAA 887

Qy 639 agactcaagaagctctccctttaataatacaagatagctgtgttatttcccca 698

Db 888 AGAACTTCA----- 896

Qy 699 gtgctactaagtgctgtctgttnttgytaagytatacacagaatccicaaagatgcaaa 758

Db 897 -----AAGTACAAACAAAAATTTAGAGAAAGCTAA 926

Qy 759 ggaatttggcataaagaagactataagctcacaagatgctctgtgtgtgtgttctt 818

Db 927 AAGAAATTTGGATTAAGAAAGCTATACGCGCCAAACATTTCTATTTGGTGGCGCTTTCTATT 986

Qy 819 tatgaatggaaccatgtagctcttctgtgtatggaaccctctgtatcttcaatgagaa 878

Db 987 GATCATATCATATATGCTGTGGCTTTCTGTATGGAGACTTCCTTGCTCTCCAGTGA 1046

Qy 879 accgtgataaccacgggaagctctgtctgttctttagtgaatccatagagatga 938

Db 1047 -----ATATACTATTGGACAGGTAAGTCACTGCTTTCTTTCTGTATTAATTGGGGCTTT 1100

Qy 939 ttgcatgtgagcagagctccctcaattgaaaccttgcgaatagcccgagagagctgctt 998

Db 1101 TAGATTGTGACAGGCAATCCCAAGCATTTGCAAAACGCAAGAGGAGCAGCTTA 1160

Qy 999 tcatatttccaggtatgtatgaagaaccagatagataacttctccacagctggagta 1058

Db 1161 TGAATTTCTTCAAGATTAATTGACATAAAACCAAGCTTTGACAGCTTTGAGAGAGTGACA 1220

Qy 1059 taacactgataccatagagaaggaactgtggaatttaaaatgttcttcaattatccalc 1118

Db 1221 TAAACCATATATATTAGGAAATTTGGAATTCAAAATGTTCACCTTACGTTACCTTC 1280

Qy 1119 aagacatctatcaagaatctggaaggtctgaatctcagaatgaatgaagctgtgagagacgt 1178

Db 1281 TCGAAAACAAAGTTAAGATCTTAAAGGTCCTCAACCTGAAGGTTCAAGAGTGGGCAACAGT 1340

Qy 1179 cgcctgtgtcgttccatagcagtgagtggaagatgagtagtgcagcttctcagaagctt 1238

Db 1341 GCGCGTGTGGGAACATGCTGCGGGAAGACACGACCTGCGAGTATGACAGAGCT 1400

Qy 1239 atatgataccagatgagcttcatcatgtgtgagatgagaatgacatcagaagctttaaagt 1298

Db 1401 CTATGACCCACAGATGCGATGCTGTATTGATGACAGACGACATTTAGGACCATTAATGT 1460

Qy 1299 ggcgcatatcagagacatattgagagtggttagtcaagaagctgttcttgcggagcaac 1358

Db 1461 AAGCATCTTGGGAAATTAATGCTGTGAGTCAAGAGCTGTGTTGCCACAC 1520

Qy 1359 catcgtaacaatatcaagatgtaagcagatgagatgtaactgtatgagatgagagagc 1418

Db 1521 GATAGCTGAAGAAACATTCCTGTGCGCGGGAATATTCACATGATGATGAGATTGAGAAAC 1580

Qy 1419 agcaagggaagcaaaatgcatatgattatcatagagatcttccataaattatatacat 1478

Db 1581 TGTATAGGAAACCAATGCTGTATGATTTATCATGAACTACTATTAATTTGACACTCT 1640

Qy 1479 ggttgggggaaaaagagctcaaatgagtgagagggcagaaacagagatcgcgaattgctcg 1538

Db 1641 GGTGTGAGAGAGAGGGCGCGGTGAGTGTGAGCAGAAACAGAAATCGCCATTGCTCG 1700

Qy 1539 tgcctttagtgaaccaccaagatctgattttagatgaagcttgcctgccttgatctt 1598

Db 1701 GGCCTGTGTCGCAACCCCAAGATTTCTGTCTGATGAGGCAACGTCAGCTCTTGACAC 1760

Qy 1599 agaaagcaagtcagctgttcaagctgtcaccttgagaagc 1636

Db 1761 TGAAGTGAAGAGAGTGTTCAGGTGCGCCCTGATTAAG 1798

RESULT 8

E02326 4378 bp RNA PAT 29-SEP-1997

LOCUS Multidrug resistance relating gene derived from human normal cells.

DEFINITION E02326

ACCESSION E02326

VERSION E02326.1 GI:2170561

KEYWORDS JP 1990100680-A/1.

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 4378)

Ueda,K. and Komano,T..

HUMAN NORMAL CELL-DERIVED MDR RELATED GENE

Patent: JP 1990100680-A 1 12-APR-1990;

SUNTORY LTD

COMMENT

OS Homo sapiens

PN JP 1990100680-A/1

PD 12-APR-1990

PE 05-OCT-1988 JP 1988251475

PI UEDA KAZUMITSU, KOMANO TORU

PC C12N15/12,C12N1/21,C12Q1/68;

CC strandedness: Single;

CC topology: Linear;

CC \*source: tissue-Type-Adrenal gland;

CC \*source: clone-SM1132;

FH Key

FT 5'UTR 1..137

FT CDS 138..3980

FT mat.peptide 138..3977

FT /gene='Multidrug resistance relating gene' FT

FT /gene='MDR1',

FT /gene='MDR1',

FT 3'UTR 3981..>4378.

FEATURES

source

1..4378

/organism='Homo sapiens'

/db\_xref='taxon:9606'

BASE COUNT 1315 a 818 c 1057 g 1188 t

ORIGIN

Query Match 25.5%; Score 514.4; DB 6; Length 4378;

Best Local Similarity 62.0%; Pred. No. 4.4e-11;

Matches 917; Conservative 0; Mismatches 471; Indels 90; Gaps 3;

Qy 162 caggtgacccctgatatgttgatagatggtgtgtgctgctgatttggttatataca 221

Db 473 CAGGTATGCTTATTTATACAGTGAATGCTGCTGGGCTGTGCTGCTTACATTTCA 532

Qy 222 gatctcctgtgatatataacacagcagcagcagcaagaagagattcgaaaacagtttt 281

Db 533 GGTTCATTTTGTGCTGCTGCGAGCTGGAAGACAAATTAACAAATTAAGAAACAGTTT 592

Qy 282 tcatcagtttggcagcagagacatcgctgtgttatagctgttgacatcgtgtgaactaa 341

Db 593 TCATCTTATTAAGCAGACAGAGATAGCGTGTGATGTGCACAGATGTTGGGAGACTTAA 652

Qy 342 cactcgcatgacagacatg---acaaaacatgtagttagttagagataaagatgctt 398

Db 653 CACCCGACTTACAGATGATGATGCTCCAAAGATTAATGAAGAAATTTGATGGAAT 712

Qy 399 gttgttcaaaacatgctactcttcttcgattgctgcctgagctgtgttgtaagagctg 458

Db 713 GTTCTTTCAGTCAATAGGCAACATTTTTCACCTGGGTTTATAGTAGATTACACGTGTTG 772

Qy 459 gaaatcaacctagtgactctatcagctctcccttataatgagcttcaagcgagcagatg 518

Db 773 GAAAGTAAACCCCTTGATGATTTTGGCATCAGTCTGTCTTGTGACATGTCAGTGTCTCG 832

OY	519	ttctagatggtcatctcctcaatgcacgtaagaaatlaagtgcctatcttccaaagcttggcg	578
Db	833	GGCAAAATATCTATCTTTCATTTTACATGATAAAGAACTTTAGGATATGGAAAAGGTGGAGC	892
OY	579	tgtgcgaagaagatcttgcatacctccgaacagctcatagctcttagggcccgagaa	638
Db	893	AGTAGCTGAAAGAGGTCTTGGCAGCAATTAGAACTGTGATGTGATTGGAGGACAAAAGA	952
OY	639	agaactccaaggctcttccctttaataataacaagatatgcttggtttatttcccca	698
Db	953	AGACTTGA-----	961
OY	699	gtggtactaatgttgttctgttcttcttgaaggtatacacagaatcccaagatgcaa	758
Db	962	-----AAGGTACAAACAAAATTTAGAGAGCTAA	991
OY	759	gaaattggcataaagaagactatagctccaagaagtgcctctgtgtctgtgtaactctt	818
Db	992	AAGATTGGGATTAAGAAAGCTTTACAGCAATATTTCTATAGTGTCTGCTTCTGCT	1051
OY	819	tatgatggaacctatggaacttcttcttggatatggaacctcttgacttctaalyaga	878
Db	1052	GATCTATGACATCTTATGCTGTGGCCTTGGATGGATGGAGCACCTTGTCCTCAGGGGA	1111
OY	879	acctggataaccatcgggaacttcttctgttcttcttcttagtgaatccatagcaatga	938
Db	1112	-----ATATTTCTATTGTGACAAAGTACTCACTGATTTCTTTCTGTATTAAATTTGGGGCTTT	1165
OY	939	tgtgacttggagcagcagctccctcacttctgaaccttcgcaatcagccgaagagctgcct	998
Db	1166	TAGTGTGTGGACAGGCATCTCAAGCATTTGACGATTTGCCAAATGCCAAGGAGCAGCTTA	1225
OY	999	tcataittcccaagttatgtataagaaccagatagatagaatacttccacagcttggata	1058
Db	1226	TGAATCTTCAAGATATATGATATATAAGCAAGCAATATTGACACTATTTGGAAGAGTGGCA	1285
OY	1059	taaacctgtaacatcagaagaactctgtgaatttaaaaatgttcttccaattcatc	1118
Db	1286	CAAAACGATATATATTTAAAGGGAATTTGGAAATTCACAAATTTTCACTTACGTTACCCATTC	1345
OY	1119	aagccatcctcaagaattcttgaagaagctcgaatccagaattaaagctctgagagaacagt	1178
Db	1346	TCGAAGAAGATTAAGATCTTTGANAAGGTCTGAACTGAAAGGTGACAGATGGGCAAGCGT	1405
OY	1179	cgacttggctcggtcactatgacatgagtggaagatcagtgtagtccacagctcttgcagaagtt	1238
Db	1406	GGCCCTGGTTGGAACAGTGGCTGTGGGAAGACACAAAGTCCAGCTGATCAGAGGCT	1465
OY	1239	atatgataccggaatgagtgttaltcatatgcttgcgaatgaatgaatcaatcagaagcttaatgtc	1298
Db	1466	CTATGATCCCCCAACAGAGGGAGTGTGCTGTGATGGACAGGATATTATAGACCATCTAAAGT	1525
OY	1299	gcggcaatcatcagagcaactatgtagtgaagttagtcaagaagccggtttgttctggagcaac	1358
Db	1526	AAGGTTTCTACGGGAAATCATTTGCTGTGTGATGATCGAAGACCTGTAATTGTTGCCACACC	1585
OY	1359	catcagtaacaatatccaagttatggaagcagatgtagtgcagatgtagaagatgtagaagc	1418
Db	1586	GATAGCTGAATAACATTTGCGTATAGGCGGTGAATAATGCACCATGATGATGAGAAAGC	1645
OY	1419	agcaagggaagcaaatgcgtatgattatcatatgtagagtttccctaataatlaatacat	1478
Db	1646	TGTCMAAGGAACCAATCTCTATGACTTTATCATGAAGAACTGCCCTCATTAATTTGACACCTT	1705
OY	1479	ggttggggaaaaaagagctcctaattgagtgtaggcaagaacacagagatctgcaattgtccg	1538
Db	1706	GATTGGAAGAGAGGGGCCAGTTGAATGGTGGGCGAAGACAGAGGATGCCATTGGACG	1765
OY	1539	tgccttagttcgaaccccaagaattctgattttagatgtaggctcaagctctgccttgatc	1598
Db	1766	TGCCCTGGTTGTGTAACCCCAAGATCCTCCTGCTGATGAGGACCAAGCTTACGCTTGGACAC	1822
OY	1599	agaaagcaagtcagctgtlccaagctgtcacttggagaag	1636

Db 1826 AGAAGCCAGACGATGCTTACAGCTGCTCTGATAGG 1863  
 ||||| | ||| ||||| | ||||| ||||  
 RESULT 9  
 LOCUS 149610  
 DEFINITION Sequence 2 from patent US 5641508.  
 ACCESSION 149610  
 VERSION 149610.1 GI:2471830  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 Unclassified.  
 AUTHORS Li, L. and Lishko, V. K.  
 TITLE Method for delivering melanin to hair follicles  
 JOURNAL Patent: US 5641508-A 2 24-JUN-1997;  
 FEATURES  
 SOURCE 1. 4646  
 /organism="unknown"  
 BASE COUNT 1371 a 892 c 1129 g 1254 t  
 ORIGIN

Query Match	25.4%	Score 512.8	DB 6	Length 4646
Best Local Similarity	62.0%	Pred. No. 1e-110		
Matches 916	Conservative	0	Mismatches 472	Indels 90
			Gaps	
Oy	162	caggttgaaccctgatacttattgcttggaataagatgtctcgtccgtgatttttggatcataca	221	
Db	760	CAGGTAATGCCATTATATTACAGTGGAAATGGTCTGGGGGCGGTTGCGCTTACATTCA	819	
Oy	222	gatttcccttggatbataatactgcagcacgacagacaagagatctcgaaacagtttt	281	
Db	820	GGTTTCATTTTGGTGGCTGGCAGCTGGAGACAAATACAAATAATGAAACAGTTTTT	879	
Oy	282	tcattcagtttggcaccagagacatcggtcgtgtttatagcttggacatcgtgaaattaa	341	
Db	880	TCATGCTTAATAATGCGACAGAGATAGGCTGTTTATGTGACACAGATGTTGGGAGCTTAA	939	
Oy	342	cactcgcatagacagaattcgac---aaatcagtgatggtlatctggagaataagatgtct	398	
Db	940	CACCCGACTTACAGATGATGTCTCTAAAGATTAATGAAGTTATTGGCGACAAATTTGAAT	999	
Oy	399	gttgtttcaaaaacatgctcacttttccgttggccggcagttgtgttggigaaggcgt	458	
Db	1000	GTTCCTTCAGTCAATGGCAACATTTTTCCTGGGTTATATGATGAGATTTACACGTGGTGG	1059	
Oy	459	gaacatcacccatagtaactctatccacgctccctcttlaaatggtctcagcgagcatg	518	
Db	1060	GAAGCTAACCCCTTGATTTTGGCCATCAGTCTCTGTTCTTGGACGTGACGTGCTGTG	1119	
Oy	519	tctctagatggtcatcctcatctgaccagtaagaattaagtcgctattccaaagctgggac	578	
Db	1120	GGCAAGAATACATCTATCTTCACTTATGAAGAACTCTTAGCGTAGCAAAAGCTGGAC	1179	
Oy	579	tgtggagaagaaggtcttctacaaatccgaacaagctcatagcctttlagggccagagaa	638	
Db	1180	AGAGTGTGAAGAGGTCTTGCGACCAATTAGAACTGTGATTTGCCATTTGGAGGACAAAGAA	1239	
Oy	639	agaactcaaaagttcttcttcctttaaatalaacaagaatagtcttggltttattttccca	698	
Db	1240	AGAACTGA-----	1248	
Oy	699	gtggctactaaagtgtgtctgtlntttgtaaggtatcacagaatctcaagaatgcaaa	758	
Db	1249	-----AAGTCAACCAAAATTTTGAAAGAAACTTA	1278	
Oy	759	ggaatttggcataaaaagacataagcttccaaaagttcctctgtgtgctgtgactctt	818	
Db	1279	AAGCAATTGGGATTAAGAAAGCTTTTACAGCCCAATATTTCTATAGGTGCTGCTTCTGCT	1338	

OY	819	tatgaatgaaacctatgtaactgtccttlttgatagaagaaaccttcattcttaataagaga	878
Db	1339	GATCATGTCATCTATTATGTCTGGCCCTTGCGTATGTGGACACACTTGGTCTCCTCAAGGGGA	1398
OY	879	aacctgatataccacccggagacgtcttcgtcgctgtttctttagtgaatcaatagacaatta	938
Db	1399	-----AATATCTATTGGACAAGTACTCACAGTATTCTTTTCTGTATTAATTGGGGGCTTTT	1452
OY	939	ttgcattggagcagcagctccctcaactttgaaacccttcgaatagcccaggagcgtgcctt	998
Db	1453	TAGTGTGTGGACAGGCATCTCCMACGATTTGAAGCAATTTGCCAAATGCAAGAAGCAGCAGCTTA	1512
OY	999	tcatattttccaggtatttgtataagaaacccaagtatactaacttttccacgctcgata	1058
Db	1513	TGAAATCTTCAAGATAATTGATTAATAAACCCAAGTATTTGACACACTTATTTGAAAGATGGCCA	1572
OY	1059	taaacctaatccatagagaagacgtgtgaaatttaaaaatgtcttcttcaatatcatac	1118
Db	1573	CAAACCACTATATATTAAAGGGAATTTGGSAATTCAGAAATGTTCACCTTCAGTTACCCTATC	1632
OY	1119	aagaccattatcaaaactcttgaaaagtcctgaatcctcagaattaaagtcctgagaagacgt	1178
Db	1633	TGCAAAAGAAGTTAGATCTTTGAAGGGCCCTGAACTGAAAGTGACAGATGGGCACAGCGT	1692
OY	1179	cgccttgcgtcggtctcctaactgacgtggtggaagaagatgcgtgtagtcacgcttctgcagaagtt	1238
Db	1693	GGCCCTGGTGTGAAACAGTGGCTGGCGGAAGGCAACACAGTCCAGCTGATGCAGAGCGCT	1752
OY	1239	ataatgccggaatgtagtgccttatcaatgbtgaabagaatgacatcacagactttaatgt	1298
Db	1753	CTATACCCCCACAGAGGGGATGGTCAAGTGTATGACAGCATTAATTAGACCATTAATGT	1812
OY	1299	gcgcattatcagagacatatgtgagttggttaagtaagaagccgtgttctgttcggagcaac	1358
Db	1813	AAGGTTTCTACGGGAATATCATTTGGTGGTAGCTCAGAGAACCTGTATGTTTGGCACCAAC	1872
OY	1359	catcagtaaccaatatcaaaglatgacagagaatgactgactgaatgaagaatgagagagc	1418
Db	1873	GATTACCTGMAAACATTCCTGCTATGGCCGCGMAAATGTCCACATGATGAGATTGACAAAAGC	1932
OY	1419	agcaagggaagaagaatgtgatatgatttatcaatlgagtttccctaataattaatacact	1478
Db	1933	TGTCAAAGGAAGCATTCCTATGACTTTATATCAATGAAACGCTCTCAATAATTTGACACCTT	1992
OY	1479	ggtaggggaaaaaaggagatcctcaatlagtgaaggaagacagaacagagatccgactgcctcy	1538
Db	1993	GCTTGAGAGACAGAGGGGCCCACTTGAATGATGTGGGACAGAACAGAGATGCCATTGGCAGC	2052
OY	1539	tgaccttagtcgaaaccocaaagattcgaattttagatgaagtagtacgtctgcgccgatttc	1598
Db	2053	TGCCCTGGTGTGCCAACCCCAAMATCCTCCTCGATGATGAGGCCAGCTGACCTTGGAGAC	2112
OY	1599	agaagcaagtcagctgtttcaacgtcgcgaactgagaag	1636
Db	2113	AGAAAGCGAAGCATGTTGCAGTGGCTCTGGATAGG	2150
<hr/>			
RESULT	10		
HUMMDR1			
LOCUS	HUMMDR1	4646 bp	mRNA
DEFINITION	Homo sapiens P-9lycoprotein (PGYL)	mRNA, complete cds.	PRI
ACCESSION	M14758		
VERSION	M14758.1	GI:187468	
KEYWORDS	P-9-lycoprotein; drug resistance protei;n; transport protein. Human drug-resistant cell line KB-C2.5 cDNA to mRNA, clones lambda HDR[10, 5, 104].		
SOURCE			
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	Chen,C.J., Chin,J.E., Ueda,K., Claik,D.P., Pastan,I.,		
AUTHORS	Gottesman,M.M. and Koninson,I.B.		
TITLE	Internal duplication and homology with bacterial transport proteins		

[illegible]

```

variation      2065
                /gene="PGY1"
                /note="c in [1]; t in [2]"
variation      3101
                /gene="PGY1"
                /note="g in [1]; t in [2] Ala-> Ser"
variation      3859
                /gene="PGY1"
                /note="c in [1]; t in [2]"
variation      4460
                /gene="PGY1"
                /note="a in [1]; g in [2]"
                /note="c in [1]; g in [2]"
BASE COUNT    1371 a      892 c      1129 g      1254 t
ORIGIN         154 bp upstream of Arai site; chromosome 7q21.1.

Query Match    25.4%; Score 512.8; DB 9; Length 4666;
Best Local Similarity 62.0%; Pred. No. 1e-110;
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;

QY 162 caggtgacccgtatattatgttggaataggtgtgtcgtccttgaattttgttacatca 221
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 760 CAGGTATGCTATTATTATACGTGGAATGTGCTGGGTGCTGCTGCTTACATTC 819
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 222 gattccttgatataactgcagcagcagcagcagcagcagcagcagcagcagcagc 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 820 GATTTCATTGTCCTGGCAGCTGAGACAAATACACAAATATGAAACAGCTTTT 879
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 282 tcaatcagtttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 880 TCATGCTATATGCGACAGAGATAGCTGTTGATGTGCACGATTTGGGAGCTTAA 939
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 342 cactcgcatacagacatgac---aaacagtgatgatatgagataagatgtctc 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 940 CACCCGACTTACAGATATGTCTCTAAGATTAAAGATTATGTCACAAAATTTGGAAT 999
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 399 gtgtttcacaacatgtctactcttctcgtatggtcgtgcagcagcagcagcagcagc 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1000 GTTCTTACAGCAATGGCAACATTTTACAGTGCTTATAGTAGATTACACGTGGTTG 1059
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 459 gaaactaaccttagtactatccacgtctccctctataatagcttccagcagcagcagc 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1060 GAACCTAACCTTGTGATTTTGGCATCAGTCTGTTCTTGACGTGTCAGTGTCTG 1119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 519 ttccagtagtgcataccatgacagcagcagcagcagcagcagcagcagcagcagcagc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1120 GGCAAAATACTATCTTCTATTTACTGATTAAGAACTCTTAGCGTATGCAAAAGCTGAGC 1179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 579 tctggcagaagaagctctgtcatccatccagacagcagcagcagcagcagcagcagc 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1180 AGTAGCTGAAGAGGCTTGGCAGCAATTAAGATGTGATTTGAGAGGAGCAAAAGAA 1239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 639 agaactcacaagcttctcccttaataatacaagaatagctgtgtttatcttcccca 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1240 AGAAGCTTGA----- 1248

QY 699 gtggctaactaagtgctgtctgtnttctgtaaggtatcacagaatcacaagaatgcaca 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1249 -----AAGTACACAAATAATTAGAAAGAAAGCTAA 1278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 759 ggaatttggcataaaagactatagcttcaaaagtgctctgtgtgtgtgtgtgtgtgt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1279 AAGAAATGGAATAAAGAAAGCTATTAACAGCAATATTTCTATAGTGTGCTTCTTCTGCT 1338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 819 tatgaatggaacctatgactgttttggataggaacctccctgattccttaatagaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1339 GATCTATGCACTTATGCTGTGGCTTGTGTATGGAGCACTGTGGCTCTCAAGGGA 1398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 879 acctgatatccatcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1399 -----AATATCTATGTGACAAAGTACGTATCTTTTCTGTATTAATTTGGGCGCTT 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 939 ttgcatltagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 1453 TAGTGTGGACAGGCAATCTCCAGCAATTGAGCATTTGCCAAATGCCAAGAGCACCTTA 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 999 tcatatttccaggttatatgataagaacccaglatagataactttccacagcagcagcagc 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1513 TGAATCTTCAATATATATGATTAATTAAGCCAAAGTATTTGACAGTATTCGAAGGTGGCA 1572
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1059 taaactgtaatccatagaagaagcagcagcagcagcagcagcagcagcagcagcagcagc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1573 CAAACAGATTAATTAAGGAAATTTGGAAATTCAGAAATTCACCTTCAGTTACCCATC 1632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1119 aagaccatcatalaagatttctgaaggtctgtaatccagaataaagtcagcagcagcagc 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1633 TCGAAAAAGATTAAAGATCTTGAAGGCTGAACCTGAAGCTCAGAGTGGCAGAGCGT 1692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1179 cgccttgctcgtctcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1693 GGCCTTGTTGGAAACAGTGTGCTGGGAGACACAAACAGTGTGATGAGAGGCT 1752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1239 atatgacggatgatgagccttcatatgctgagtcagtcagtcagtcagtcagtcagtc 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1753 CTATGACCCACAGAGGAGTGTGATGATGACAGGATATTAAGACATTAATGT 1812
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1299 ggggcatatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1813 AAGGTTTCTACGGGAAATCAATGTGTGTGTGAGTCAAGAACCTGTATTTGCCACAC 1872
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1359 catcagtaacatacacaatcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1873 GATAGCTGAAACATTTGGCTATGCGCTGAAATGTCACCATGATGATGATGAAAGC 1932
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1419 agcaagggaagcaaatgcgatgatttcatcagtcagtcagtcagtcagtcagtcagtc 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1933 TGTCAAGGAAGCCCAATGCTTATGACTTATCATGAACTCCCTATTAATTTGACACCT 1992
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1479 ggttaggggaaaaagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1993 GCTTGAAGAGAGAGGCGCCAGTTGAGTGTGTGAGTCAAGAACAGCAGAGATCCCATTTGCACG 2052
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1539 tgccttagtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2053 TGCCCTGTTGCAACCCCAAGATCTCTGCTGATGAGAGGCCAGCAGCAGCTTGACAC 2112
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1599 agaaagcaagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2113 AGAAAGCAAGCACTGTGTTCAAGTGTGCTGCTGATTAAGG 2150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 11
AR091275 LOCUS AR091275 4669 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5994088.
ACCESSION AR091275
VERSION AR091275.1 GI:10018030
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4669)
AUTHORS Mechetner,E. and Roninson,I.B.
TITLE Methods and reagents for preparing and using immunological agents
specific for P-glycoprotein
JOURNAL Patent: US 5994088-A 1 30-NOV-1999;
FEATURES Location/Qualifiers
source 1..4669
BASE COUNT 1393 a 894 c 1130 g 1252 t
ORIGIN

```

```

Query Match    25.4%; Score 512.8; DB 6; Length 4666;
Best Local Similarity 62.0%; Pred. No. 1e-110;
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;

```

OY	162	caggttgcacccctgattatgatttgaaatagggtgtctgccttgatttltggttaacata	221
Db	760	CAGGTATGCCCTATTATTATACAGTGGAAATTGGTGTGGGGTGGTGGTTCGCTTACATTTCA	819
OY	222	gattcccttgtagatlaactgcagacagacagacagagatttcgaaacagtttt	281
Db	820	GGTTTCAATTTTGGGCTGGCAGCTGGAGACAAATACCAAAATTTAGAAAACAGTTT	879
OY	282	tcattcagtttgtagacagagacacgcgttgattgtagctgtgaacgtgtgaactaa	341
Db	880	TCATGCTATTATGGACAGAGATAGGCTGGTTGATGTGCACGATGTTGGGACCTTAA	939
OY	342	cactgcacagacagacatg--acaaatcagtgatggtatctgagataaagtgtctc	398
Db	940	CACCCGCACTTACAGATGATGTCCTCCAAAGATTATGAAAGAAATTTGGTGCAAAATTTGGAAT	999
OY	399	gttgtttcaaaaacatgctcacttttcgatttgccgcgagttgttggtggaaggtcg	458
Db	1000	GTGTTTTCAGTCATATGGCAACATTTTTCACGTGGTTTATATGAGGATTTTACACGTGTGG	1059
OY	459	gaacatccctatgtagctatccacagctcctccttaataatgtagctcagcgacagat	518
Db	1060	GAACTTAACCTTGTGATTTTGGCCATTCAGTCTGTTCTTGACATGTCAAGCTGCTGTG	1119
OY	519	tctcagaatgtagcatcattgacccagtaagaaatgaatgaatgaccttccaaacgtgagc	578
Db	1120	GGCAAAATACTATCTTCAATTTTACTGATTAAGAACTTATAGGATTTGCAAAAGCTGAGC	1179
OY	579	tgtgacgaagaagatctctgtgcataatccgaaatgcataagccttaaggccagagaa	638
Db	1180	AGTGTGTAAGAGGTCTTGGCACCAATTATGAACTGTGATTTGATTTGGAGCAAAAGAA	1239
OY	639	agaactccaagaagtccttccctttaaataaacaagatagcttgattatltcccca	698
Db	1240	AGAACTTGA-----	1248
OY	699	gtggtactaagtgtgtctgttntlttgaaatatacacagaatctcaaatgacaa	758
Db	1249	-----AAGGTACACAAAATTTNAGAAAGCTAA	1278
OY	759	ggatttcggcataaaaagacatagcttcaaaagtctctctgtggtgtgtactctc	818
Db	1279	AAGATTGGGATTAAGAAACATTTACAGCAATATTTCTATAGGTCGTCTTCTCTGCT	1338
OY	819	taigaatggaacctagacttccttttgtagtaagaaacctctgattcttaatgagga	878
Db	1339	GATCTATGACATCTTATGCTGTGGCTTGTGTGTGGACACCTGTGTCTCTCGAGGGA	1398
OY	879	acctggtataccaatcgagactgttctgcgtgttcttcttagtgaatccatagagta	938
Db	1399	-----ATATTCTATTTGGACAAATACTCACTGATTTCTTTCTGTATTAATTGGGGCTTT	1453
OY	939	ttgcatgtgagcaagcagctccctcacttctgaaacctgcgaatagccagagagctcct	998
Db	1453	TAGTGTGGACAGCACTCCACAGCAATTTGAGCAATTTGCAAAATGCAAGAGGACAGCTTA	1512
OY	999	tcaatattccaggttatctgataagaaccagtatagataactttccacagctcgata	1058
Db	1513	TGGAATTTTAAAGTAATTTATATATAGCCAACTATTTGACACTATTTGCAAAAGTGGGCA	1572
OY	1059	taaacctggaatccatagaagaactgtggaatttaaaatgtcttccatataatccatc	1118
Db	1573	CAAAACGATTAATTTAAGGGAATTTGGAAATTCAGAAATGTTCACATTGATTAACCAATC	1632
OY	1119	aagaccatatacaagaattctgaaagtctgaaatctcagaatlaaagctctgagaaagct	1178
Db	1633	TCGAAAAGAAAGTTAAGTCTTTGAAGGGCCGAAACCTGAAAGTGCAGAGTGGGACAGAGCT	1692
OY	1179	cgctctggtgcgtctcaatgacgtgagtggaagatacgttaagctcagctctgcagaagt	1238
Db	1693	GGCCCTGGTTTGGAAACAGTGGCTGTGGGAAGACCAACAGTCCAGCTGATCCAGAGCT	1753

QY	1239	atatatccggaatgttgcttattcaatggtgagatgacatcaagctttaaagt	1298
Db	1753	CTATGACCCACAGAGGGGATGTCAGTGTGATGACAGATATTTAGGACATTAATGT	1812
QY	1299	gggagatcatcgaagccatatgtgagttgtagtcaagaagccgtgtcttctcgggaacc	1358
Db	1813	AAGGTTTCTACGGGGAATTCATTGTTGGTGTGATGACGAACTGTATTTGTCACACAC	1872
QY	1359	catcgtatacaatcaatcaagatcgaagatgtagtgcgtgaagaagatgagagaagc	1418
Db	1873	GATFACCTGAAAACATTCGCTATGCGCCGGAATAATGTACACATGATGAGATGAGAAAGC	1932
QY	1419	agcaagagggaagcaaatgcgtatgattatcaatcgtgagtttcttaataatcaacat	1478
Db	1933	TGTCAAGGAAAGCAATGCGCTATGACCTTATCATGTAAAGTGCCTCATTAATTTGACACCT	1992
QY	1479	ggtaggggaaaaagagctcaaatgaagtggaggggaagaagaagaagatcgaattgccc	1538
Db	1993	GGTGGAGAGAGAGGGGCCCACTTGAAGTGGTGGGAGAAAGCAGAGATGCGCATTTGACG	2052
QY	1539	tgccttaagtgaagaccaccaagatctgattttaagatagagtcgtccctgattc	1598
Db	2053	TGCCCTGTGTGCCAACCCCAAGATCTCTCTGCTGATGATGAGCCACGCTTGACAC	2112
QY	1599	agaaagcaagtcagctgttcaagctgcacatcgagaag	1636
Db	2113	AGAAAGCAAGCAGTGGTTACAGTGGCTGTGATTAAG	2150
RESULT	12		
LOCUS	108557	4669 bp	PAT
DEFINITION	Sequence 3 from Patent WO 8705943.		02-DEC-1994
ACCESSION	108557		
VERSION	108557.1	GI:588735	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 4669)		
AUTHORS	Robinson,I.B., Pastan,I.H. and Gottesman,M.M.		
TITLE	COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS		
JOURNAL	Patent: WO 8705943-A 3 08-OCT-1987;		
FEATURES	location/Qualifiers		
source	1..4669	/organism="unknown"	
BASE COUNT	1394 a 892 c 1129 g 1254 t		
ORIGIN			
Query Match	25.4%;	Score 512.8;	DB 6; Length 4669;
Best Local Similarity	62.0%;	Pred. No. 1e-110;	
Matches 916; Conservative	0;	Mismatches 472;	Indels 90; Gaps 3;
QY	162	caggttgaccccgattatgttggaatagtggttcgctcgatttttgftacataca	221
Db	760	CAGGATGCGCTATTTATTCAGTGGAAATGTGCTGGGGTGGTGGCTGCTTCAATTCOA	819
QY	222	gatttcctgtgattataatacgaagcagcagcaagaagattcgaanaacgtttc	281
Db	820	GGTTTCATTTTGTGTCGCCGACGCGGAAGACAATTAACAACAATTTAGAAAAAGATTMTT	879
QY	282	tcaatcagtttggcacagagacatcgctggtttgtagatagctgtgacatcggtgaactaa	341
Db	880	TCATGCTTAAATGACAGAGAGATAGCGCTGTTGATGTGACGAGATGTTGGGGAGCTTAA	939
QY	342	caatcgatgaagaagcatgac---aaatcgtaatgtagtatggaagtaagatggcct	398
Db	940	CACCCGACTTACAGATGATGCTCTTAAGATTATTAAGTAAGTTATGTGTGACAAAATTGGAA	999
QY	399	gttgtttcaaaacgtctactcttttcgattgcgcctcgagctgtgttggtgaagggctg	458





```
D6 2641 -----AAGTACACAAAAATTAGAAAGCTAA 2670
OY 759 gatttggcataaaaggactatagctcaaaagtgtctctgtgtgtactctt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2671 AAGAAATTGGGATTAAGAAAGCATTTACGCCAATATTTATAGGCGCTTCTGCT 2730
OY 819 tatgaatgaaactatgactgtcttctgtcatgaaacctcttcttaattgaga 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2731 GATCTATGATTTATGCTCTGCTCTGCTATGGAGACCACCTTGCTCTCAAGGGA 2790
OY 879 acctgataaccatcggaactgtcttctgtcttcttcttgaatccatagaagta 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2791 -----ATATCTATGGACAGTACTCAGTATCTTTCTTGATATATGGGCTTT 2844
OY 939 ttgcattgagagcagcttccctcaactttgaaaccttcgaatagcccgagagctgcct 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2845 TAGGTGTGGACAGGATCTCCAAAGCATTTGAAACATTTTGCAGAAAGAGAGACACTTA 2904
OY 999 tcatatttccaggttatgtatagaagaaccagiatagataacttctccacagctgata 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2905 TCAATCTTCAAGATTAATTGATTAATAGCCAGTATTCACACTTATTCAGAGTGGGA 2964
OY 1059 taaccctgaatccatagaagaagactgtgaaattaaatgttcttcaattatccalc 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2965 CAACCCAGATTAATTAAGGAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCTATC 3024
OY 1119 aagaccatcatcaaatgttcaaaagctgtgaatcagaatgaatgaatcgtgagagagct 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3025 TCGAAAGAGTAAAGTAAATCTTGAAGGCTGAACCTGAAGGTCAGAGTGGGACAGCTGT 3084
OY 1179 cgccttgcgtgtccatcagctgtgagagagctgtgagagctgtgagagagct 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3085 GGCCCTGCTGGAACAGAGGCTGTGGAAGAGCACAAAGTCCAGCTATTCAGAGGCT 3144
OY 1239 atatgattccgagatgtgtcttcatatggtgtgagaaatgaatcagaagctttaatgt 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3145 CTATGACCCACAGAGGGAGTGTGATGTGATGAGCAGATTAATTAATGAT 3204
OY 1299 ggggcattatcgaagacatattgtgagtgatgaagagcctgttcttgcgagacac 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3205 AAGGTTTCTACGGGAATCATTTGCTGTGAGTCAAGAACTGTAATTTGCCACAC 3264
OY 1359 catcagtaacaatcaaatgtatgagcagagatgtgagactgtgagagatgagagagc 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3265 GATTAGCTGAACCAATTCCTATGAGCCGTGAATAATGTCACATGATGATGAGAAAC 3324
OY 1419 agcaaggagaaagcaatgtcgtatgatttatacagtgagtttccataaattatcatc 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3325 TGTCAAGGAAGCCATGCTATGACTTTTTCATGAAACGCTCATTAATTTGACACCT 3384
OY 1479 ggttaggggaaagagctcaaaatgagtgagggcagaagaaacagagatgcgaattgtcg 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3385 GGTGTGAGAGAGAGGCGCCAGCTTGAGTGTGTCGAGAGAGAGATGCGCATTTGACAG 3444
OY 1539 tgccttgaatgcgaaccccaagaattctgattttagatgagagctgctgtccctgattc 1598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3445 TCCCTGCTGCTCCAAACCCCAAGATCTCTGCTGATGAGGCCACGCTTGACAC 3504
OY 1599 agaaagcaagtcagctgttcaagtgtcagctgtcagctgtgagaag 1636
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 3505 AGAAAGCAAGCAGTGTTCAGGTGCTCTGGATTAAG 3542
```

```
RESULT 14
AX012321
LOCUS AX012321 8630 bp DNA PAT 06-SEP-2000
DEFINITION Sequence 2 from Patent EP0955374.
ACCESSION AX012321
VERSION AX012321.1 GI:9998370
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
          artificial sequence.
```

```
REFERENCE 1 (bases 1 to 8630)
AUTHORS Baum,C.D., Hildinger,M. and Osterlag,W.P.
TITLE Retroviral vectors for gene transfer
JOURNAL Patent: EP 0955374-A 2 10-NOV-1999;
HEINRICH PETTE INST (DE)
FEATURES
    source
        1..8630
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        /note="Proviral Plasmid-DNA"
    misc_feature
        1..160
        /note="Plasmid-R ckgat (pUC)"
    misc_feature
        1..8630
        /note="Retroviral Expressionsvektor SPbeta1MSA1"
    misc_feature
        161..677
        /note="5'-LTR"
    5'UTR
        532..1219
        mat_peptide
            1220..5062
            /note="MSA1 mdr1 CDNA"
    misc_feature
        5215..5774
        /note="3'-LTR"
    misc_feature
        5775..8630
        /note="Plasmid-R ckgat (pUC)"
BASE COUNT 2234 a 2011 c 2189 g 2196 t
ORIGIN
Query Match 25.4%; Score 512.8; DB 6; Length 8630;
Best Local Similarity 62.0%; Pred. No. 1.1e-110;
Matches 916; Conservative 0; Mismatches 472; Indels 90; Gaps 3;
OY 162 caggttgaccctgattatgattggaatgattgttgcctgatttggtaactaa 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1555 CAGGTATGCCCTATTATTAAGTGAATGTGCTGGGCTGCTGCTTACATTTA 1614
OY 222 gatttctgtgtatatactgacagacagacagacagagagatcgaacagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1615 GGTTCATTTTGTGCTGCTGCGACCTGGAAGACAAATACAAATTAAGAAACGTTTT 1674
OY 282 tcaatcagtttggcagagacatcgctgtgttgaatgactgtgacatcggtgaactaa 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1675 TCATGCTAATATGCCAGAGATAGTAGTGTGATGATGACAGATTTGGGAGCTTAA 1734
OY 342 cactgcagtagacgaattgac-----aaatcagtgatgtatgtgagataaagattcct 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1735 CACCCGACTTAAGAGATGATGCTCTTAAGATTAAAGATTATGAGCAAAATTTGGAAT 1794
OY 399 gttgttcaaaacaattctacttcttgcattgagcctgtgcagttgtgtgagagctg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1795 GTTCTTTCAGTCAATGCGCAACATTTTTCACGTGGTTTAATGATTAACAGTGTG 1854
OY 459 gaaactcacctagtgactctatccacgctctcctctataatgtgttcagcgagcagctg 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1855 GAAGCTAACCTTTGATTTTGGCATCATGCTGTTCTTGACACTGACCTGCTG 1914
OY 519 ttctagatggtcatctcatcttgcacagtaagataaagtctcatccaagctgggc 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1915 GGCAGAAAGATCATCTTTCATTTAAGTAAGAACTTAAAGCTTAAGCTAAGCAAAACCTGAGC 1974
OY 579 ttgtgagagaagatcttctcatcaatccgaacagcatagcctttagggcagaagaa 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 1975 AGTAGCTGAAGAGCTTTGGACACATTAAGAACTGATTTGAGAGCAAAAGAA 2034
OY 639 agaacttcaaaagttcttccctttaaataaacaagatagctgtgttatttcccca 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2035 AGAAGCTTGA----- 2043
OY 699 gttggtactaaagtgtgtctgttntttaggtatatacagaagatctcaagatgcaaa 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D6 2044 -----AAGGTCAACAAATAATTTAGAAAGAGCTAA 2073
OY 759 gatttggcataaaaggactatagctcaaaagtgtcttgtgtgtgtactctt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db	2074	AAGAAATTGGGATTAAGAAACCACTTTACAGCCAAATATTCTCTAAGSTGCTGCTTTCCTGCT	2133
Qy	819	tatgaatcgaaactatagacttgcttcttggtatcgaaaccttccttgattcttaatygaga	878
Db	2134	GATCTATGCAATCTTATGCTCTGCTGCTTCTGATATGGACACCACTTGCTCTCCAGGGGA	21939
Qy	879	accctggatatacatcctggagactgtcttgctgttcttcttagtctgaatcaatagagta	938
Db	2194	-----ATATTCTATTGGACAATGACTACTGATATTCTTCTGTATTAAATTTGGGGCTTT	2247
Qy	939	tgtcatgtgagcagcagctccctcaacttctgaacctgcgcgaatagaccgagagctgcctt	998
Db	2248	TAGTGTGGACAGGCACTCCACAGCAATTGAGATTTGCAAAATGCAAGAGGACAGCTTA	2307
Qy	999	tcatatttccaggttatgtataagaaccgatatagataacttctccacagctygata	1058
Db	2308	TGAATCTTACAGTATATTGTATTAAGCCAAATATTACACTTATTCGAAGAGAGGCGCA	2367
Qy	1059	taaacctgaaatcatatgaagaagacttggaaatttaaaaatgttcttcaattatcatc	1118
Db	2368	CAAAACCGAATTAATTAAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCATTC	2427
Qy	1119	aagaccatcatcaagaatttcgaaagcttcgaatcccaagatlaagtcttggagagagct	1178
Db	2428	TGCAAAAAGAAAGTTAAGATCTTGAAGGGCCTGAACCTCGAAGGTGCGAGAGTGGCAGACGT	2487
Qy	1179	cgccttgctgcgtctcaatlygcaglyggaagatagcgttagctccagctcttcgcagagtt	1238
Db	2488	GGCCCTGTTGGAACACAGTGGCTGTGGGAAGCACAAAGTCCAGCTGATGCAAGAGCT	2547
Qy	1239	atatgatccgagatgagtgcttcatgtgtgatgtagaatgacatcagagctttaatgt	1298
Db	2548	CTATGACCCCAACAGAGGGAGTGGCTGATGACAGGATATTAGACCATTAAGT	2607
Qy	1299	gcggcatatccgaagacatatgtgagttgttagcaagaagcctgttcttgctggagcaac	1358
Db	2608	AAGTTCCTACGGGAATCTTGTGTGTGTGACTCAGCACTGTATTGTTCACCAAC	2667
Qy	1359	catcagtaacaataatcaagatagacgagatgatatgtacatgataagaagatlgagaagac	1418
Db	2668	GATAGCTGAAAAACATTCGTATGGCCGTGAAATGTCAACCATGAGTGAAGATTGAGAAGC	2727
Qy	1419	agcaagaagcaaatgctgatatattatcatatgtaggttctctaataattataact	1478
Db	2728	TGTCAAAGGAAGCCATGCCATTAACCTTTACAGAAACGCCCTCAAAATTTGACACCT	2787
Qy	1479	gttgggggaaaaagagctcaaatgtatgtagggcgagaaacagagatcgcgaattgctgcg	1538
Db	2788	GGTTCGAGAGAGAGAGGSCCAAGTTGAGTGTGGGCGAGAAAGCAGAGATGCCATTGGACG	2847
Qy	1539	tgccttagttcgaaaccccaagaattctgattttagatgtagcctagcgtctgccttgattc	1598
Db	2848	TGCCCTGTTGCGCAACCCCAAGATTCCTCGTGTGATGAGGCGACGTCAGCCTTGGACAC	2907
Qy	1599	agaaagcaagtcaagctgttctaagctgcgaactgagaag	1636
Db	2908	AGAAAGCGAAGCAAGTGGTTCAGGTGCTGCTCGATTAGG	2945
RESULT	15		
AR028672			
LOCUS	AR028672	9318 bp	DNA
DEFINITION	Sequence	6	from patent US 5858744.
ACCESSION	AR028672		
VERSION	AR028672.1	GI:5940645	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 9318)		
AUTHORS	Baum,C., Stocking-Habbers,C. and Osterstag,W.		
TITLE	Retroviral vector hybrids and the use thereof for gene transfer		
JOURNAL	Patent: US 5858744-A 6 12-JAN-1999;		

FEATURES	Location/Qualifiers
Source	1..9318 /organism="unknown"
BASE COUNT	2377 a 2217 c 2364 g 2358 t 2 others
ORIGIN	
Query Match	25.4%; Score 512.8; DB 6; Length 9318;
Best Local Similarity	62.0%; Pred. No. 1,1e-110;
Matches 916; Conservative	0; Mismatches 472; Indels 90; Gaps
Oy 162	cagcttgaccctgattatgttgaaatagtgtgtgccttgatttttggttacaata 221
Db 2111	CAGGTATGCCATTATTATTCAGGTGGAATTGGTGTGGGGTGTGGTGTCTTACATTCA 2170
Oy 222	gatttccttgtgatttlaactgcaagcagacagacaaagagatttcgaaacagtttt 281
Db 2171	GTTTCAATTGTTGGTGTGGCAGCTGGAGACAAATATACAAATATTAGAAAACAGTTT 2230
Oy 282	tcaatcagttttgacagagacatcggtcgtgtttgatagcctgtgacatcsgtgaactaa 341
Db 2231	TCAATGCTATATATGCGAGAGAGATAGCGTGTGTTGATTCGACGATGTTGGGGAGCTTAA 2290
Oy 342	cactcgcgatgacagacattac---aaatcagtgatgtatgtgagaataagattgcct 398
Db 2291	CACCGGACTTACAGATGATGTTCTCTAAGATTATATGAGATTATTTGGTGACAAATATGGCAT 2350
Oy 399	gtgttttcaaaaacatgtcctatcttttcagattgacctgcgcagttgtttcgttgaaagcgt 458
Db 2351	GTTCTTTACGTCATATGGCAACATTTTTCACGTGGCTTATATAGTAGATTATACACGTGTG 2410
Oy 459	gaacatcaccttagtgactctatccacgctccctcttlaatgactcagcgagcatg 518
Db 2411	GAAGCTTAACCCCTTGATGTTTGGCCATCATGATCCTGTCTTGACATGTCAGTGTGTCTG 2470
Oy 519	ttctcagatgttcatctcattctgaccagtaagaattaaigtctcttccaagctcgtgggc 578
Db 2471	GGCAAAAGTACTATCTTCATTTTCTGATTAAGAAACCTTAGGGTATGCAAAAGCTGGAGC 2530
Oy 579	tgtgcagaagaagctcttgcacatccagcaagcagcatagccttlaagcccgaaagaa 638
Db 2531	AGTAGCTGAAGAGGTCTTGGCAGACANTTAGAACGTGATTTGGATTGGAGGACAAAGAA 2590
Oy 639	agaacttcaaaagtccttccctttaaatabaacaagatatgctcgtgtttatlttccca 698
Db 2591	AGAACTTGA----- 2599
Oy 699	gttgctactaatgctgtgtctgttntttgaaggtatlaacgaagatccaaagatgcata 758
Db 2600	-----AAGGTCAACAACAAAATTTTGAAGAGAGCTTAA 2629
Oy 759	ggaatttgcataaaaagacataagcttcaaaagtcctcttggtgtgttactctt 818
Db 2630	AAGATTGTGGATTAAGAAAGACTTATACAGCCAAATTTTCTATAGGTGCTGCTTCTCTCT 2689
Oy 819	tatgaatggaacctatgacctgtcttttggatlaagaaacctccatgattcttaatlygaa 878
Db 2690	GATCATCATCTTATGCTGCTGGCGCTTGTGATGGAGCACCTGGTCTCTCAGGGGA 2749
Oy 879	acctgatatcacgctcggaactgttcttgcgttttcttctttagtgaatccatagcagtta 938
Db 2750	-----ATATCTTATTTGGACAAGATACTCACTGTATTTCTTTTGTATTAATTGGGGCTTT 2803
Oy 939	ttgatattgagcagcagctccctcaactttaaacttcgaaactcgaatcgaagcgagagctgcct 998
Db 2804	TATGTTTGGAGAGGCACTTCCAAAGCATTTGAAGCATTTGGAAATGCAAGGGAGCAAGCTTAA 2863
Oy 999	tcaatatttccaggcttatatgataaagaaccagatataaacttttccacagcctggtga 1058
Db 2864	TGAATCTTCAAGATATATGTATAATAAGCAAGATATTTGACACTTATTTGGAAGAGTGGGCA 2923
Oy 1059	taaacctgaatccatagaagaacgttggaaattttaaagatgttcttccaattatccatc 1118

```
Db 2924 CAAACCAATATATTAAGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCAATC 2983
QY 1119 aagaccatcatcaagatctcgaaggtcgtgaatctcagaattaagtcggaagacagt 1178
Db 2984 TCGAATAAGAGTTAAGATCTTGAAAGGCGCTGAACCTGAAGGTGACAGAGTGGCAGACGGT 3043
QY 1179 cgcccttgctcggtcgaatgagcagtgaggaagagtagcaggttagtcagcttcgcaaggtt 1238
Db 3044 GGCCCTGGTTGGAACACAGTGGCTGTGGGAAGAGCACACACAGTCCAGCTGATGCAGAGGCT 3103
QY 1239 atatgaccggaatgctggtctatcatggttgatgagaatgacatcagagctttaatgt 1298
Db 3104 CTATGACCCACACAGAGGGAGATGCTCAGTTGATGACAGATATTAGACCATTAATGT 3163
QY 1299 gcgcatcatcgaagaccatattggaagtgttagtcaagagcctgtttgttcgggaccac 1358
Db 3164 AAGGTTTCTACGGGAATCATTTGTTGTGTGAGTCAGGAACCTGTATTGTTTGCACCCAC 3223
QY 1359 catcagtaacaatatcaagtatgacgagatgctgactgatagaagatgagagagc 1418
Db 3224 GATACCTGAAAAATTCGCTATGCGCCGTGAATAATGTACACATGATGAGATTGAGAAAGC 3283
QY 1419 agcaaggaagcaaatgctgatalctatcatggaagttccctaaataataataacatt 1478
Db 3284 TGTCAAGGAAGCCAAATGCTTATGATGAATACTGCTCATTAATTGACACCCT 3343
QY 1479 ggttaggggaaaaagagctcaaatgagtggagggcagaagaagagatcgcaattgctcg 1538
Db 3344 GTTTGAGAGAGAGGGGCCCACTTGAGTGTGGCAGAGAGGATCGCCATTGACAG 3403
QY 1539 tgccttagctcgaacccccaagaatctgattttagatgaggtacgctcgcctgagatc 1598
Db 3404 TGCCTGTTTCACAAACCCCAAGATCTCTGCTGAGTGAAGGCCACAGCTCAGCTTGACAC 3463
QY 1599 agaaagcaagtcagctgttcaagctgcaactgagagaag 1636
Db 3464 AGAAAGCGAAGCAGTGTTCAGTGTGCTCTGATAAG 3501
```

Search completed: April 22, 2002, 22:55:04  
Job time: 27812 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 22:55:04 ; Search time 14771.4 Seconds

(without alignments)  
2166.654 Million cell updates/sec

Title: US-09-873-409-16

Perfect score: 1940

Sequence: 1 ttccgcttgcgcgatgact.....taaatgcctacttataga 1940

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2344280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_hlgo\_hum:\*  
31: em\_hlgo\_inv:\*  
32: em\_hlgo\_rod:\*  
33: em\_hlgo\_hum:\*  
34: em\_hlgo\_inv:\*  
35: em\_hlgo\_rod:\*  
36: em\_hlgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	624.6	32.2	4279	6	AX105082 Sequence
2	623	32.1	4279	6	AX105078 Sequence
3	623	32.1	4279	6	AX105080 Sequence
4	621.4	32.0	4045	12	AF269224 Synthetic
5	621.4	32.0	4279	6	AX105057 Sequence
6	619.8	31.9	4317	4	AF045016 Canis fam
7	619.8	31.9	4317	6	AX105059 Sequence
8	605.4	31.2	4378	6	E02326 Multidrug r
9	603.8	31.1	4646	9	HUMMDR1
10	603.8	31.1	4669	6	AR091275 Sequence
11	603.8	31.1	4669	6	I08557 Sequence 3
12	603.8	31.1	6505	6	AR028671 Sequence
13	603.8	31.1	8630	6	AX012321 Sequence
14	603.8	31.1	9318	6	AR028672 Sequence
15	603.8	31.1	8630	6	AX012320 Sequence
16	602.2	31.0	8630	6	AX012320 Sequence
17	597.4	30.8	4669	6	AR055785 Sequence
18	595.8	30.7	4186	6	AX108654 Sequence
19	595.8	30.7	4195	6	AX108656 Sequence
20	594.4	30.6	4264	6	AR051647 Sequence
21	592.8	30.6	4264	6	AR051650 Sequence
22	589.4	30.4	3988	6	AX024454 Sequence
23	589.4	30.4	4192	9	AF015535 Sequence
24	581.2	30.0	4280	10	CR0PGP11
25	579.8	29.9	4489	4	AB029153 Sequence
26	579.6	29.9	4296	10	CR0PGP11
27	578	29.8	3987	10	CR0PGP1185
28	578	29.8	4304	10	CR0PGP1165
29	571.6	29.5	4298	10	MUSMDR
30	568.6	29.3	3858	4	OU78609
31	568.4	29.3	4233	6	AR123273 Sequence
32	568.4	29.3	4356	10	MUSMDRAA
33	568.4	29.3	4924	10	MUSMDR1A
34	566.8	29.2	4254	10	RATMDRM
35	562	29.0	4018	5	GGA9799
36	560.2	28.9	4323	10	AF286167 Rattus no
37	560.2	28.9	4927	10	AF257746 Rattus no
38	557.4	28.7	4390	5	XLU17608 Xenopus lae
39	536.4	27.6	3924	6	AX024455 Sequence
40	536.4	27.6	3924	9	HUMMDR3
41	528.4	27.2	3912	10	RATPGLYCO
42	528.4	27.2	4084	10	MUSMDRA
43	523.6	27.0	3905	10	CR0PGP11
44	503.6	26.0	3682	10	MUSPGLYIA
45	441	22.7	4280	9	AF136523 Homo sapi

## ALIGNMENTS

RESULT 1  
AX105082  
LOCUS AX105082 4279 bp DNA  
DEFINITION Sequence 26 from Patent WO0123540.  
ACCESSION AX105082  
VERSION AX105082.1 GI:13921232  
KEYWORDS  
SOURCE  
ORGANISM  
dog.  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 4279)  
Patten,C.J., Steimel-Crespi,D.T., Crespi,C.L., Reiff,T.C. and  
Stocker,P.J., 1999. Canis familiaris and uses thereof  
Patent: WO 0123540-A 26 05-APR-2001;  
GENTEST CORPORATION (US)  
LOCATION/Qualifiers  
1. 4279  
/organism="Canis familiaris"

CDS

```

/db_xref="taxon:9615"
17..3862
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC37723.1"
/db_xref="GI:13921233"
/translation="MDPEGGRGSAEKNFMKMGKSKKKEKKKPTVSTFAMFRYSN
WDRILVGLTMAAIIHGALPLMLVFNMTDSFNAGISNRKTPVLIINEITNT
OHFINLHEEMTYAYVYSGIGAVLAAYIOVSEFMAAGROILIKROFPAIKRO
EIGMPFVHVGELNPTDQVSKINTEGIDSKIMFOSIATFPTGGTIVGTBMRKIL
VILASPVGLSAIWAIRKILTSPTDKELIAYAKAGAEVLAIAPIYALVSEVIGK
EERKNLEBAKGIKIKATITANISGAELIYASALAPWYTSVLSEVIGKYL
IVFESVLGAESIGQASPSIEAFANRGAAYELFKIIDKPSIDSYSKGKHNDKNG
NLEKFNHVSYSRKEVKTILKGLNKVSGOTVALNGSGCSTVOLMORLYDPTD
GWCIDGODIRITINVRHEITGVSOEPLFATTIENIRYREVMWDEIKAYKE
AANADEVIMLPKPNKDFLVGERGKOLSGOKRIAIRALVIRNKKIILLDEANSALDE
SEAVYDALDKARKGRTTYIAHRLSTVNRADVIAGFDGVIYEKRNHDEMEKRTY
FLVIMOTIGNEIELENATGESKESDLEMSKDSILKTRSTRSRIHAQGDOR
KLGTEDELENNVPVPSFWRLIKLNTSEWPEVVGICAIINGLOAFSIFSRIGI
FRDEDEPERRKONSMEVFLVGIISFTFLOGFTFGKAGEILTKRLRYWFRSM
LRDVSFMDPKMTGATLTLRLANDAAOVKAGISRLAYITOMIANGIGIISLIG
MOTLILLIIVPITAIAGVEMKMLSGALKKKELEAGKIAETAEIENFRYVSTFR
EOKFEYMAOSLOVPRNSLRKRIHIFGVSFTIOAMMYTSYACFPFGATLVANEMN
FODVLVESAIIVGAAVAVOVSSFAVDYAKAKVSAHVIMITEKSLIDSYSPHLKP
NLEGVNTEVEVFNPTRPDIPLVGLGLEVKKQTLVGVSGGKSTVVOJLBRF
YPLEGASVLIDGKEIKHLNWOMLRHGIIVSOEPIIFDCEIENIAYGNSRVSHEE
YVOAKENAIHRIETPLPERKYNRVGDKTOLSGGOKRIAIRALVIRNKKIILLDEA
TSALDESEKVOEALDKARREGRTCIIVAHRLSTIQNADLIIVFQNGKVEKHTHQL
LAQGIYFVMSVQAGAKR"
BASE COUNT      1296 a      833 c      1009 g      1141 t
ORIGIN

```

```

Query Match      32.2%; Score 624.6; DB 6; Length 4279;
Best Local Similarity 66.5%; Pred. No. 7,6e-141;
Matches 929; Conservative 0; Mismatches 459; Indels 9; Gaps 2;

QY 162 caggtgacccbtatattgttggaaatggtgtgtccttgatttggtaacata 221
    || || || || || || || || || || || || || || || || || || ||
DB 358 CAGGTAGGCTATTATTACAGTGGAGTGGCTGCTGCTGCTGCTGCTGCTACATCA 417

QY 222 gattctcttgatataactcagcagcagcagcagcagcagcagcagcagcagc 281
    || || || || || || || || || || || || || || || || || || ||
DB 418 GGTTCATTTGTGGCTGGCAGCAGAAAGATCTCAAAATTTGAAAACAAATTTT 477

QY 282 tcaatcagtttgagcagcagcagcagcagcagcagcagcagcagcagcagc 341
    || || || || || || || || || || || || || || || || || || ||
DB 478 TCATGCTATATGCGACAGAGATGGCTGGTTTGACGTGCATGACCTTGGGAGCTTAA 537

QY 342 cactcgacatgacagacatg--acaaatcagtgatglatgtagaataagatgctc 398
    || || || || || || || || || || || || || || || || || || ||
DB 538 CACCCGGCTCAGAGAGATGCTCCAAATCAATGAAGGATTGGCACAATAATTGGAAT 597

QY 399 gttgttcaaaacatgtctactcttctcagatgacctgagcagtggttggtaaggc 458
    || || || || || || || || || || || || || || || || || || ||
DB 598 GTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTATGTGGGGGTTTACACGTGGTTG 657

QY 459 gaacacacactagtagctatccacgctctcctcttaataagctcagcgagcagatg 518
    || || || || || || || || || || || || || || || || || || ||
DB 658 GAAGCTAACCTTGTGATTTTGGCCATCAGCCTGTTCTTGAGCTTTCAGCCGCTTCTG 717

QY 519 ttctagatggtcatcctatgacagtaagaaatgaagctcattccaagctggggc 578
    || || || || || || || || || || || || || || || || || || ||
DB 718 GGCAGAAATACATCTTACTTACTGATTAAGAACTCTTGSCGATCAAAAGGTGAGC 777

QY 579 tctggcagaagaagctctgtcacaatcgaaacagtcataagcctttagggcccaaggaa 638
    || || || || || || || || || || || || || || || || || || ||
DB 778 AGTAGCTGAAGAAGTCTTACACCAATCAAGACTGTGATTTGCTTTGGAGACCAAGAA 837

QY 639 agaacctcaaggtatcacacagaatcacaagatgcaagatttggcctaaagaagac 698
    || || || || || || || || || || || || || || || || || || ||
DB 838 AGAAGCTTGAAGAGTACACAAAATTTAGAGAAGCTTAAAGGAATTGGGATTAAGAAAGC 897

```

```

QY 699 tatagctcaaaagtgctctcttggtctgtgtactcttcttatgaatgaagcaatagact 758
    || || || || || || || || || || || || || || || || || || ||
DB 898 TATCAGCGCAACATTTCTATTGTTGGTCCGCTTCTTATTGATCATCATCATATGCTCT 957

QY 759 tgccttttgatagaacccctctgattctcttaatgagaacccggtatataccatcgagc 818
    || || || || || || || || || || || || || || || || || || ||
DB 958 GCGTTTCTGTATGGAGCTCCCTGGTCTCTCCAG-----TCAATATATCTATTGGACA 1011

QY 819 tttctgctgcttctctcttgatgaatccataagctatgactgaagaacgagcc 878
    || || || || || || || || || || || || || || || || || || ||
DB 1012 AGTACTACACTCTCTCTTCTTCTGATTAATTTGGGCTTTAGTATTTGACAGGATCC 1071

QY 879 tcaatgaaaccttcgaatagccagaggagctgaccttcaatatttccaggtatga 938
    || || || || || || || || || || || || || || || || || || ||
DB 1072 AAGCATTTGAAGCATTTGCAACGCAAGAGGAGCGCTTATGAATCTTCAAGATTAATGA 1131

QY 939 taagaacacagtatagataacttccacagctgtagataaacctgaaatccatagaagg 998
    || || || || || || || || || || || || || || || || || || ||
DB 1132 CAATTAACCAAGCATTTACAGCTATTGGAAGAGTGACATAAACCATATATTAAGGG 1191

QY 999 aactgtgaatttaaaatgttcttctcaattatccataagaccatctcaagatctc 1058
    || || || || || || || || || || || || || || || || || || ||
DB 1192 AAATTTGGAATTCAAAATGTTCACTTCACTTACCTTTCGAAAAGAACTTAAGATCTT 1251

QY 1059 gaaagctgtaatctcagaatlaagctcggagagacagctgccttggtcgtctcaa 1118
    || || || || || || || || || || || || || || || || || || ||
DB 1252 AAAGGCTCTCAACTGTAAGGTTACAGATGGGACAGACAGTGCGGCTGTTGGAACACTGG 1311

QY 1119 caatgggaagaagctagctgtagtccagctctcggagagttatagatccgagatgagctc 1178
    || || || || || || || || || || || || || || || || || || ||
DB 1312 CTGCGGGAAGGACGACGACCGGCTGATGACAGGCTCTATGACCCACAGATGGCAT 1371

QY 1179 tatcatgtgtagatagaatcatcagatcagctttaaactgtcgcatcatcagaacat 1238
    || || || || || || || || || || || || || || || || || || ||
DB 1372 GGTCTGATTTGATGACAGACAGCATTAGACCATTAAGTAAGCATTTGGGAAATTTAC 1431

QY 1239 tgaagtggttagtcaagaagcctgttctgtcggagaccacatcagtaacaatalcaagta 1298
    || || || || || || || || || || || || || || || || || || ||
DB 1432 TCGTGTGTGATGACGAGAGCGCTGTGTGTTGCCACACATATGCTGAAAACATTTCCGTA 1491

QY 1299 tgaacgagatgagtgtgactgtagaagatgtagaagcagcagcaggaagaacaatgctga 1358
    || || || || || || || || || || || || || || || || || || ||
DB 1492 TGGCGGCAAAATGTCACATGATGAGATTGACAAAGCTTTAAGAACCCATTTGCTTA 1551

QY 1359 tgaattatcatgaggttctcctaataattatatacttgaagggaaagaagagctca 1418
    || || || || || || || || || || || || || || || || || || ||
DB 1552 TGATTTTATCATGAAGAACTACTTAATTAATTTGACACTCTGTTTGACAGAGAGGGCCCA 1611

QY 1419 aatgagtgagaaggcagaagaagagatcgcaatgtgctgtgacctagttcgaaaacccaa 1478
    || || || || || || || || || || || || || || || || || || ||
DB 1612 GCTGAGTGTGAGACAGAAAGAAATCGCCATTCGTGGGCGCTGTTGCAACCCCA 1671

QY 1479 gatctgattttgataagctacgctcgtcccttgagatcagaagaagcagctgtca 1538
    || || || || || || || || || || || || || || || || || || ||
DB 1672 GATTCCTTCTCTGTGATGAGCAACGTCAGCTGTGACACTGAAAGTAGAACAGTGTTC 1731

QY 1539 agctgcactggaagaag 1555
    || || || || || || || || || || || || || || || || || || ||
DB 1732 GGTGGCCCTGTGATRAGG 1748

```

```

RESULT 2
AX105078
LOCUS AX105078 4279 bp DNA
DEFINITION Sequence 22 from Patent WO0123540.
ACCESSION AX105078
VERSION AX105078.1 GI:13921228
KEYWORDS
SOURCE
ORGANISM dog.
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE
1 (bases 1 to 4279)

```





[illegible]

OY	459	gaacatcccttgatctaacatccacgctccctcttlaatgacctgaagcgagacatg	518
Db	658	GAACCTAACCCCTGTGTATTTTGGCCATCAAGCCCTGTCTTGAGACTTTCAGCGGCATCTG	717
OY	519	tctaggatgtaatcctcattgacccagtaagaaatlaagtgctattccaaagcttgggc	578
Db	718	GGCAAAAGATCTATCTTCAATTTACTGATTAAGAAACCTTTGGCTATGCAAAAGCTGGAGC	777
OY	579	tgtggcagaagaagctctgtgcatcaatccgaacagctcaatagacgttttagggcccgagaa	638
Db	778	AGTAGCTGAAGAAACTCTTACACCAATTCAGAACTGTGATGGCTTTGGAGACAAAAGAA	837
OY	639	agaactccaagatcatcaagaatctccaagatccaagaatttggcatcaaaaagagc	698
Db	838	AGAACTTGAAAGGTACACAAAATAATTAGAGAGCTAAAGGAATTTGGATTAAGGAAGAC	897
OY	699	tatagctcaaaaagtgtcctctgtgtgcgtgtgactctttalgaatgtaaacctatgact	758
Db	898	TATCACGGCCAACTTCTATTTGGTCCCTCTTTGATGATCTATGCATCATGTGCTCT	957
OY	759	tgtcttttgatggaaccccttggtttcttaatgagaactgtaataccacggagc	818
Db	958	GGCTTTCTGGATGTGGACCTCTTGGTCTCTCCAA-----TGAATTTCTATTGGACA	1011
OY	819	tgtcttgagcttttctttagtgtaatccaagcaatgcatcttggagcagagctcc	878
Db	1012	AGTACTACTGTCTCTTTCTTTCTGTATTAAATGGGGCTTTTACTATTGGACAGGCATCCCC	1071
OY	879	tcaacttgaacacttcgaataagccgagagagctgtcccttcatalttccagttatgta	938
Db	1072	AAGCATTAACCATTTTGCAAAAGCAAGAGAGAGAGCTTATGAATCTTCAAGATTAATTGA	1131
OY	939	taagaacaccagatataactttccacagctgtgatatlaaactgtatccatagaag	998
Db	1132	CAATTAACCAAGCATTTACAGCTATTTCGAAGCTGACATTAACCAAGATTAATATTAAAGG	1191
OY	999	aacgttggaattttaaataatgttcttccaattatccataaagaccatctacaagattct	1058
Db	1132	AAATTTGGAATTTCAAAAATGTCTACTTCACTTACCTTACCTTTCCAAAAGAATTAAAGTCTT	1251
OY	1059	gaagctgcgatctccagaaatlaagcttgtagagagaagctgcgctgtgtcgtatcctaag	1118
Db	1252	AAAGGCTTCAACCTGGAAGGTCAAGAGTGGCGCAGACAGTGGCGGTGGTGGGAACAGTGG	1311
OY	1119	cagtggaagaagtaacgtatgtaacgcttctgcgaaggttatatgatacggatgatagctt	1178
Db	1312	CTGCGGGAAGGACAGACCGTGACGTGATGAGAGAGCTCTATGACCCACACAGATGGCAT	1371
OY	1179	tatactgtgatagagaatgaacatcaagagtttaagtgtggagcatlatacgagacatat	1238
Db	1372	GGTCTGTATTAATGAGACAGCAATTAAGACATTAATGTAAAGCATCTTGGGAAATATTAC	1431
OY	1239	tggatgtttagtcaagaagcgttltgttgcggaaccacatcagtaacaatalcaagta	1298
Db	1432	TGCTGTGTTAGTCAAGGAGCGCTGTGTTTGGCACACGATAGCTGAAAACATTCGCTA	1491
OY	1299	tggacagagaatgatactgataatgaagaatgtagagagacgaagaaggaaagcaatatgcta	1358
Db	1492	TGCGCGCAAAATGTCAACATGATGATGATGAAGAAAGCTGTTAAGGAAGCCAAAGCTCTA	1551
OY	1359	tgaattatcatgagulttctctaataattataacttgaattgttgaaggaaagagagctca	1418
Db	1552	TGATTTTTCATGAAACTACTAATTAATTAATTTGACCTGTGGTTGGAGAGAGGGGCCCA	1611
OY	1419	aatgagtggagggcagaacaagaagatcgcgaattgtctgtgcttgaatgctgcgaaccoca	1478
Db	1612	GCTGAGGTGTGACAGAAACAGACAAATCGCATTTGCTCGGGCCCTGTGTTCCGAACCCAA	1671
OY	1479	gattctgattttagatgagcgtactgcttgccttggatltcaagaagcaagtcagctgtca	1538
Db	1672	GATTTCTTGTGCTGATGATGAGCAAGTCGACGCTGGAACACTGAAAGTAAGTGAAGTGGTTCA	1731
OY	1539	agctgcactggaagaag	1555

Db 1732 GGTGGCCCTGATAAGG 1748

RESULT	4			
LOCUS	AF269224			
DEFINITION	AF269224	4045 bp	DNA	SYN
ACCESSION	AF269224			
VERSION	AF269224.1	GI:8926216		

CDS

BASE COUNT	1171 a	850 c	986 g	1038 t
ORIGIN				

Query Match	32.0%	Score 621.4;	DB 12;	Length 4045;
Best Local Similarly	66.4%;	Pred. No. 4.5e-140;		
Matches 927; Conservative	0;	Mismatches 461;	Indels 9;	Gaps 2

QY 162 caggttgacctgatatatgttggaaagagtgttcgtgccttgatttttggtaacataa 221  
 |||||  
 Db 460 CACGTATCCCTATTATTACAGTGGATGGGCTGGCTGCTGCTCTTACATCCA 519  
 |||||  
 QY 222 gatttccttgtgatatctaactcagcagcagcagcagaagagatttgcgaacacagtttt 281  
 |||||

Db	520	GGTTATTCTCGTGCGCTGGCAGCAGGAAAGACAGATACATCAAAATTAGAAAAACATTTT	579
OY	282	tcattcagtttggcacaagaacatcgcgtgtgtttagtagctgtgtcacatcgtgtgaacttaa	341
Db	580	TCATGCTATCATGTCGAGACAGAGATATGGCTGTTCACGTGCATGACGTTGGGGACCTTTA	639
OY	342	cactcgcgatgacagacattt---acaaatcagtgatgtatgttgaataagatgtcct	398
Db	640	CACCCGGCTCACAGACCATTTCTCCAAAATCAATGAAAGAAATTTGGCGACAAAATTGGCAT	699
OY	399	gttgtttcaaaaacatgtcctacttcttcgatattggtccgtggcagttgtgtgtgaagagctg	458
Db	700	GTTCTTTCAATCAATACACAACTTTTTCACCGCTTTATATAGTGGGTTTACACGCGGTG	759
OY	459	gaactacacctagtgactctatccacgctccctcttataatgtccttaagcggcagcatg	518
Db	760	GAACCTAACCCCTGTGTGATTTTGGCCATCAGACCCCTGTTTGAGCTTTCAGCCGCATGTG	819
OY	519	ttctagaatgttcatctcattgacccagtaagaattaaigtccctattccaaagctcgggc	578
Db	820	GGCAAAATCTATCTTCAATTTACTGATTAAGAAACCTTGGCCTATGCAAAAGCTGGAGC	879
OY	579	tgtgacgaagaagcttctgtcatcaatccgaagaagcatcagctctttagggcccgagagaa	638
Db	880	AGTAGCTGAAGAGTCTTATGACCAATCAGAACTGTGATGTGCTTTGGAGCACAAGAA	939
OY	639	agaacttcaaaagfatcacagaatccaaagatgtcaaaagatttlygcatataaaagac	698
Db	940	AGAACTTGAAGAAGTACACAAAATTTTGAAGAGCTTAAAGAAATTTGGGATTTAAACAAAGC	999
OY	699	tatagcttcaaaagtgtctctgtgtgtgtgtgttcttcttataatgtgaagactatgact	758
Db	1000	TATCACGGCCAAACTTTCTATTTGGTGGCGCTTCTTATGATCTATGATCATATATGCTCT	1059
OY	759	tgcttttgtgtatggaacccctcttgattctctaatsgagaacctgtgatataccatgtggac	818
Db	1060	GGCTTTCTGGATGAGGACCTCTTGCGCTCTCCAA-----TGAATATCTATTGGACA	1113
OY	819	tgctctgtcgttttcttcttagtgaatccaaagattatgcatatgtagcagcagatccc	878
Db	1114	AGTACTCAGCTCTCTTTCTTTCTGATTAATTTGGGGCTTTTAGATTTGACACAGCATTC	1177
OY	879	tcactttgaacccctgcgaatagccgagagagctgccttcatatlttccaggttatgta	938
Db	1174	AAGCATTTGAAGCATTTTCCAAACCAAGAGAGCAGCTTATGAAATCTTCAATATATTGA	1233
OY	939	taagaaccccgatatactaacttcttccacagctgtgatatlaaacctgaatccatatagaag	998
Db	1234	CAATTAACCAAGCTTTCACGCTTATGGAAGTGAAGTGAATTAACCAAGATATATTAAGG	1293
OY	999	aactgtgaacttaaaaatgttcttcttcaatcttccatacaagaacctatcaagatctc	1058
Db	1234	AAATTTGGAAATTCAAAAATGTTCACTTCAGTTACCTTCTCGAAAAGAAAGTTAACAATCT	1355
OY	1059	gaagagctgaatcctcgaatlaagtcgtgagagacagtcgccttlygtctcctaayg	1118
Db	1354	AAAGGCTCTCAACCTGGAAGTTCAAGATGAGTGGCAGACAGTGGCGCTGTGGAGAACATGG	1413
OY	1119	caatgtagaagaagtacggttagtccaagctctcgcagaggttatatagatcggatagtgctt	1178
Db	1414	CTGGGGGAAGGACAGACGCTGTGACGTGAAGTGAAGAGCTTTATGACCCACAGATGGCAT	1477
OY	1179	tatcatgtgtgaatgaatatacatcaagagtttaagtgtgcgcattatcgcgaacatat	1238
Db	1474	GCTCTGATTTGATGACAGGACATTTAGGACCAATTAATGTAAGCATCTTTGGGAAATTAC	1533
OY	1239	tggaagtgtttagtcaaaagagctgttltgttcyggacaccatccagtaacatatcaatga	1296
Db	1534	TGGTGTGGTGAAGTCAAGGAGCTGTGTTTGGCACACAGATAGCTGAAACATTCGCTA	1592
OY	1299	tgaacgagatgatgtgcctatgtgaagagatgtagagagacgacgaagggaagcaatgtcgtta	1355
Db	1594	TGGCGCGCAAAATCTCACATGTGATGATTTGAAAGCTGTTTAAGGAAGCCCAATGCGCTA	1653

QY	1339	tgattttacatgtaggttctctaataatcatatcattggttaggggaaagaagctca	1418
Db	1654	tgattttatcattgaaactacttaataaattttgacacttctggttgagaggaagggccca	1713
QY	1419	aatgagtgaaagcagaacaacagagatgcgaattgctcgtgcttagtctgaaacccca	1478
Db	1714	gctagtgctgtagacagaaacagaaatgcccattgctctggccctcctggtcccaacc	1773
QY	1479	gattctgatttagtagagagctcagctcgtccctgattcaagaagaacagctacgttca	1538
Db	1774	gattcttctgctgtagagcgaacgtcagctctgtagacactgaagtaagcagctgttca	1833
QY	1539	agctgcactgagaagaag	1555
Db	1834	ggtggccctgagataag	1850
RESULT	5		
LOCUS	AX105057	4279 bp	DNA
DEFINITION	Sequence 1 from Patent WO0123540.	PAT	30-Apr-2001
ACCESSION	AX105057		
VERSION	AX105057.1	GI:13921209	
KEYWORDS			
SOURCE	dog.		
ORGANISM	Canis familiaris		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
AUTHORS	Stocker, P.J., Steimel-Crespi, D.T., Crespi, C.L., Reif, T.C. and Paten, C.J.		
TITLE	P-glycoproteins and uses thereof		
JOURNAL	Patent: WO 0123540-A 1 05-Apr-2001; CORPORATION (US)		
FEATURES	location/Qualifiers		
source	1..4279		
CDS	/organism="Canis familiaris"		
	/db_xref="taxon:9615"		
	17..3862		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC37718.1"		
	/db_xref="GI:13921210"		
	/translation="MDPEGGRKSGAEKFNFMKGRKSKNEKKRPTVSTFAMRYSNWLDRLMYLVGTMAAIHIGCALPLMLVFGNMTDSFANAGISRNKTPVLIINESITNNPTQEHFNHLEEMTYYAAYSGIGAGVLAAAYOVFPCGLAAGROLIKRKOPFAHMQEIGFHDVHGEELNTRLLTDVSKINEGIGKGMFHSIATFFGVLVGFGRKLLLVIIASPYVLGSAAIMAKILSSFPDKELIATAKGAIAEVLAIIRVIAAGGRKLEERKVNLEAKGIGIKKAITIANISIGAFELITASVIAEYISLVLSSTYSIGVLTVPFSFISIGAFSIQASPSIEAFANAGAAIEIKIIDNKSIDSYSKSGKPNKINGNLEFRNVHFSYSPSRKEVKILGLNMLKVSQGVTAIVLNGSGGKSTVQLMQRLPDNDGMVCIIDQDRTITVNRHLRETTGVSGSPVLEATTIANIENGENTMDIEKAVEKEANAVADITMKLPNKEDTLVGERGAOLSGOALIKRLEAGAKIATEALIENTPTVSLTRSEAVVOALDKARKGRRTIYIAHLSYRNADYIAGDDGYIYKGNHDEMRKKGITYIFKLTWOTRGTEIELENNATGESKESDLEKSPDSGLIKRSTRSHAPQGDURKLGKREDLNEKRVNPFVSWRLIKLNSITMPFVVGIFCAIINGGLQAPFSIISRIIGITFTQREDEPTEKRONSNMFSVLEVLGIISITFFELQGFPGKAGELIKRLRYVYRSLMLROVSWMEDDKNTTGALTTRLANDAAOVKALISRLAVIOTNANLGTGIIISLIYGKMOILLAIIVPIIPIAIAAGVEMKLSGOALKRLEAGAKIATEALIENTPTVSLTRFOKEFYMAQSLQYPRNLSLKAHIEGVSFSTITQMMFTSAGCFRGATIVANEPMMKFQDVLVFSALVFSGMAVGVSVFAPDYAKKVAHAHIMTEKSPILDSYSPGLKRNNTLEKRVNENVEVNYPTRPDIPIVQIGISLEVKGGOTLALVSSGGCKSTVQMLLEFRFYDPLAGSVLIDGKEIKHLNOMWLRAHGIYSQEPILPDCIAENIAYGDNRSVSHSEERIMOAKEANIHHTLELPKYNTRVSGKGTLSGGOKORIAIARALVROPHIILLDEETSAIDTSEKVKYQALDKARKGRCTIYIAHLSITQANDLIVYQNGKVKENGTHQDLAQGIFISMVQAGAR"		
BASE COUNT	1294 a	834 c	1008 g
ORIGIN	1143 t		
Query Match	32.0%	Score 621.4;	DB 6;
Best Local Similarity	66.4%;	Pred. No. 4.5e-140;	Length 4279;

Matches	92% ; Conservative	0 ; Mismatches	46% ; Indels	9% ; Gaps	2
QY	162	caggtgacccctgatactgattgtgaaatagtggtgcgccttgattttgttacaata	221		
Db	358	CACGTATGCGCTATTATTAACAGTGGATCGGTCGCGGCGCTGCTGGTCTGTTACATCA	417		
QY	222	gattcccttgtagtataaactgagccgcagacccaaggagtttcgaaaacgtttt	281		
Db	418	GGTTTCATTGGGGGCCGCGACGAGAGACAGTACTCAAAATTATGAAAACATTTTTT	477		
QY	282	tcattcagtttttggocacagacatcgcgtctgttatagtcgtgtacatcgttgaactaa	341		
Db	478	TCATGCTATCATGGGACACAGAGATGGCTGGTTGGATTGACGTGCATGACGTTGGGAGACTTAA	537		
QY	342	cactcgcaatgacagacattg---acaaaatcagatgagtattggagataagattgctct	398		
Db	538	CACCGCGCTCACAGACGATGTCTCCAAATATGAATGAAGAAATTGGCGCAAAATTGGAAAT	597		
QY	399	gtgttttcaaaaacatgctacttttttgcattgtgcctgcgtgacgttgattgttgggaaggtcg	458		
Db	598	GTTCCTTTCATCAATAGACACATTTTTTACCGGTTTATAGTGGGGTTTAAACGTGGTTG	657		
QY	459	gaaactaacccctagtgactctacacagctccctctctataatgagcttcaagcgcagcatg	518		
Db	658	GAGGTATACCTTGTTGATTTTGGCATCAGCGCTGTTCTTGAGCTTTCAGCGGCATCTG	717		
QY	519	ttctagatggtgcatactcatttgaccagtaagaaatlaagtgcctattccaaagctgggac	578		
Db	718	GGCAAAAGTACTATCTTCATTCTTACTGATTAAGAAACCTTGCCCTATGCAAAAAGCTGGAGC	777		
QY	579	gttggcagaagaagctctgtcatcaatccgcagacagtcagctttaggcgcagaa	638		
Db	778	AGTACTCTGAAGAAGCTTTAGACGACATAGAACTGTGATTGGCTTTGAGACGCAAAAADA	837		
QY	639	agaaactcaaaagtatcacacagaatctcaaaagatgcaaaaggatttggcacaanaaaggac	698		
Db	838	AGAACTTGAAGAGTTCACACAAAAAATTAGAAAGACTTAAAGGAATTGGGATTAAGAAAGC	897		
QY	699	tatagctcaaaagatgctctctgtgtgtgtgtaactcttataatgaaagacctaagact	758		
Db	898	TATCAGCGCCCAACATTTCTATTGGCGCGCTTCTTATTGATCTATGACATATGCTCT	957		
QY	759	tgctttttgtatgaaacctccttgattcttctaaggaagaaactgatactacatcgggac	818		
Db	958	GGCTTTTGGTATGGGACCTCTCTGGCTCTCCAG-----TGAATATTCATTTGGACA	1011		
QY	819	tgctctgtctgtttctcttagtgaatccatagacagttatgacttggagcagcagctcc	878		
Db	1012	AGTACTACTGCTCTTCTTTCGTATTAATTGGGCTTTTAAATTTAGGACAGGCACTGCC	1071		
QY	879	tcaacttgaacacttcgcaataagcccgagagctgccttcaalatttccaggtatttga	938		
Db	1072	AAGCATTTAAGACATTTGGAAAAGGAGAGGACACTTATGAAATCTTCAAGATTAATTGA	1131		
QY	939	taagaaaccccgatagataacttttccaacgctfgatataaactgatacocatagaag	998		
Db	1132	CAATTAACCAAGCATTTGACAGCTATTTCGAAGAAGTGGACATTAACCAAGATTAATTAAAGG	1191		
QY	999	aactgtgaatttaaaaatgtttcttccaatatcatcaagaacacatcatcaaatctc	1058		
Db	1192	AAATTTGGATTTCAAAAAATGTTCACTTCACTTACGTTACCTTCTCGCAAAAAGAAAGTTAAATCTT	1251		
QY	1059	gaaagcttgatcaatcagaatlaagtctctggaagacagtcgccttgctgctgcataatg	1118		
Db	1252	AAAGGCTCTCAACCTGAAGGTTCAAGATGGGGACAGAGGGCGCTGGTGGGAACAGTGG	1311		
QY	1119	cagtggaagaagatcagtgtagtccagctcttcgagaagttatataatccgagatgagctt	1178		
Db	1312	CTGCGGGAAGACAGACAGCCGTGCACCTCATGTGACAGAGGCTCATGTGACCCACAGATGGCAT	1371		
QY	1179	tatactgtgtagtgaatgacatgaagagctttaaattgcyggaattatcgagacatat	1238		
Db	1372	GCTCTGTATTTGGAGACAGGACATTAGGACATTAATGTAAAGCAATCTTGGGGAATTTAC	1431		

Query Match	31.9%	Score 619.8	DB 4	Length 4317
Best Local Similarity	66.3%	Pred. No. 1.1e-139		
Matches 926	Conservative 0	Mismatches 462	Indels 9	Gaps 2
162	caggttgaccctgtaattatgttgaataagggtgtgtcgtccctgtaatttgytllacataca	221		
408	CACGATGCGCTATTATTACAGTGGGATGCGTCTGGCGTGCCTGGTGGCTTACATCA	467		
222	gatttccttgytgattataactgtcagcagcacagacagaccagagattcgaagaactttc	281		
468	GCTTTCATTCTTGCTGCTCCCTGGCAGCAGGAAGACAGATTCTCAAAATTAGAAAACTATTTT	527		
282	tcattcagtttggagacaggacatcggcgtgtttatagctgtgtgacatcgtgtgaactta	341		
528	TGATGCTATCATCAGCAGCAGAGATTGGCTGGTTTGACGTGCATGACGTTGGGAGCTTAA	587		
342	caactgcacatcagaacatg---acaaaaacagltatgttatcttgagataaagattgctct	398		
588	CACCGCGCTCACAGAGAGATGCTCCAAATCATAGAGAAATTGGCGCAAAAGTTGGAAAT	647		
399	gttgttcaaaaacatgtcacttcttcgatgtgctgtcagcaglttggttggagaggtcg	458		
648	GTTCTTTCAATCAATAGCAACATTTTTCACCGTTTTAATGAGGGGTTTACACTCGTTG	707		
459	gaactaccctcaotgacctataccagctctcctcttataaagtcgtgaagcgagacatg	518		
708	GAGGCTAACCCCTTGATTTTGGCAATCAGCCCTGTCTTGACATTTTCAGCCGCATCTG	767		
519	tctcagtagtgcatactatcttgcagtaaggaaatlaagtcctatcccaaaactggggc	578		
768	GGCAAGATACATCTCTTCTTACTACGATTAAGAACCTTTGGCCTATGCAAAAGCTGGAC	827		
579	tgtygcagagaagaagctctgtgtcaatcccgaaagtcatalagcctttagggccagagaa	638		
828	ACTAGCTGAGAAGCTCTTAGCAGCAATCAGAACTGTGATTTGCCCTTTGGAGCAAAAGAA	887		
639	agaactccaaggtatacacagaatcccaagaatgcaaggaatttggcatabaaagac	698		
888	AGAACTTGAAAGGTACACAAAAAATTTTAAACAAAGCTTAAACAAATTTGGATTAAGAAAC	947		
699	tatagcttcaaaaagtgctctgt	758		
948	TATCAGCGGCAAAATTTCTATTGTGATTAATTTGGGGCTTTTATGATTTGATCATTATGCTCT	1007		
759	tgcttlttgatagtgaaacctcctgtgatttcaatgtgagaaacctggtatatatacatcgagac	818		
1008	GGCTTCTGTGTAGGGAAGCTCTTGCTGCTCCAG-----TGAAATATCATATTGGACA	1061		
819	tgcttctgtcttcttctttagtgaatccatagcagttatgtcatgttgagcagtgccc	878		
1062	GGTACTCATGCTCTCTTTCTTGTGATTAATTTGGGGCTTTTATGATTTGAGCAGGCAATCCC	1121		
879	tcactttgaaccttcgaaatagcccgagagagctgccttcacatatttcacagatttga	938		
1122	AAGCATTTGAGCAATTTTGCAACGCAAGAGGAGAGCTTTTGAAATCTTAAAGATTAATTA	1181		
939	taagaagaccagtatagataaactttccacagctggtatataaactgtaatccataagaag	998		
1182	CATATAACCAACGATTTGACAGCTATTGCAAGAAGTGACACATTAACGAGTAATTAATTAAGG	1241		
999	aactgtggaatttaaaaatgttcttccaattatccatcaagacatctatacaagattc	1058		

Db	1242	AAATTGCAATTCAAAAATGTTCACTTCACTGATTACCTTCTCGAANAAGATTAAATCTT	1301
Oy	1059	gaaaggtctgcatctcagaatlaagtctcgagagacagtcgctctgctcgtctcaatg	1118
Db	1302	AAAGGCTTCAACCTCGAAGGTTCAAGAGTGGGCGAGCAATGGCGCTGTGGGACAGTGG	1361
Oy	1119	cagtcggaagagtaagcgtgaagccagctcttcagaggtatatagtatccgtagtcgctt	1178
Db	1362	CTGGGGAAGACACGACCGCTGCACCTGATGCAGAGGCTTATGACCCACAGATGGCAT	1421
Oy	1179	tatcatgttgagtagaaatgaatcagaagcctttaagtctgcgcattatcgaagacat	1238
Db	1422	GGTCTGATTGATGGACAGGACATTAGGACCATTAATGTAAAGCATCTTCCGGAATTAAC	1481
Oy	1239	tggagtgctgtaagcaagagcgtcttcttctcgagccacatcagtaacaatatagta	1298
Db	1482	TGCTGTGATGATCGAGAGCTGTGTGTTGCCACCGATAGCTGAAAACATTCGCTGA	1541
Oy	1299	tggacgagatgattgactgattgaagagatvgagagacgaaggaagcaaatgcyta	1358
Db	1542	TGGCGCGCAAAATGTCACCATGAGTGAATTTGAGAAACCTGTTAAGGAAGCCAAATGCTTA	1601
Oy	1359	tgattttcatgaggtcttccataaatttaatacatcttggtagagggaanaagagagctca	1418
Db	1602	TGATTTTATCATGAACACTAATTAATTTTGACACTCTGTTGGAGAGAGGGGCCCG	1661
Oy	1419	aatgagtgagaggcagaacaagagagatgcgaattgctcgtgcctagatcgaaacccca	1478
Db	1662	GCTAGTGTGGACGAAACAGAGAAATGCGCATTTCTCGGGCCCTGTGTTGGCAACCCCA	1721
Oy	1479	gattctgattttagatgaggtcgaagtcgtcccttgattatcagaagaagtcagctgtcca	1538
Db	1722	GATTCTTCTGCTGGATGAGGCAACCTCAGACTCTGAGACCTGAANAATGAAGCATGTTTCA	1781
Oy	1539	agctgcactgaggaagg	1555
Db	1782	GTGGCCCTGGATTAAAG	1798

RESULT	7			
AX105059				
LOCUS				
DEFINITION	AX105059	4317 bp	DNA	PAT 30-APR-2001
ACCESSION	AX105059			
VERSION	AX105059			
KEYWORDS				
SOURCE				
ORGANISM	dog.			
REFERENCE	Canis familiaris			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
TITLE	1 (bases 1 to 4317)			
JOURNAL	Stockert,P.J., Stelmel-Crespi,D.T., Crespi,C.L., Reif,T.C. and Patten,C.J.			
FEATURES	P-glycoproteins and uses thereof			
Source	Patent: WO 0123540-A 3 05-APR-2001;			
	GENTEST CORPORATION (US)			
	Location/Qualifiers			
	1..4317			
	/organism="Canis familiaris"			
	/db_xref="taxon:9615"			
	70..3912			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC37719.1"			
	/db_xref="GI:13921212"			
	/translation="MDPEGGRGSAEKNFMWKKSKKEKKEKPVVSTFAMFRYSMM			
	LDRLYMGVTMAAIIHGAALEPLMLVFGNMDSEFANAGISNNKTFPVIIINISITNNIC			
	LFHINHEEEMTAYAYVIGAGVLAIVAYIDVSEWCLAAGQAIKIRQFHPAIMROCF			
	IGFMDVDVGELEINRLFDVSKINEGIDKVGMEFOSIATPEGTIVGTFPGMKLLTI			
	ILASIPVLGSAATWARTLISFTPKELLAYAKAGVAEEVLAARIYVAFPGGCKKELE			
	RYNNLEEARKIGIKRAITANISIGAAFLIITYASALAFWGTISVLSSETTIGCVLT			
	VFEFVLIGAFSIGASPTEAFANARGAAVEIFPIIDNRPSIDSYSKGHPDNIKGG			
	LEFNHVSFSPSRKREVIKLKLNKLVOSGQVVALVGNSGCKSTTVOLMRLYDPTDC			

[illegible]

```

Db 1122 AAGCATTTGAGCATTTGGCAAGCAAGAGAGAGAGCTTATGAATCTTCAGATTAATGCA 1181
QY 939 taagaagaccagatagataaactttccacagcttgatataaacttgatccatgaag 998
Db 1182 CAAATAACCAACCATGAGACAGTATCGAAGAGATGACATAAACAGATTAATTAAGAG 1241
QY 999 aacttggaaatataaagttcttcttcaatcatcatcacaacatcacaatcattct 1058
Db 1242 AATATTGGCAATTCAAAATATGTTCACTTCACTTACCTTCTCGAANAAGATTAGATCTT 1301
QY 1059 gaagaagctcgaatcgaatgaatgaatcgtggaagacagtcgcttgctgctcgaatg 1118
Db 1302 AAAGGCTTCAACCTGAAGCTTCAGAGTGGGAGACAGTGGGCTGTGGGAGACAGTGG 1361
QY 1119 cagtggaagagatgaagtgatcagctcctcgaagaggtatataatgatcgatgatgctt 1178
Db 1362 CTGCGGGAAGACAGACAGCTGACCTGATGCAAGAGCTCTATGACCCACAGATGCGCAT 1421
QY 1179 tatcatgtgtgagatgaatgaatcgaatcgaatgaatgtgagacatcgaatcgaatc 1238
Db 1422 GGTCTGTATGATGACAGAGAGATTAAGACATTAATGAAGGCTTCTGCGGAATTAAC 1481
QY 1239 tggagtgatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1298
Db 1482 TGGTGTGTGATGATGACAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1541
QY 1299 tggagagagatgtgtgacatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1358
Db 1542 TGGCGGCAAGAAATGTCACCATGATGATGATGATGATGATGATGATGATGATGATGAT 1601
QY 1359 tgaatttcaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1418
Db 1602 TCATTTTATCAATGAACATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1661
QY 1419 aatgagtgagagagagagagagagagagagagagagagagagagagagagagagag 1478
Db 1662 GCTGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1721
QY 1479 gatcttgatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1538
Db 1722 GATTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1781
QY 1539 aactgacagagagagag 1555
Db 1782 GGTGGCCCTGATTAAGC 1798

```

RESULT 8  
 LOCUS E02326 4378 bp RNA PAT 29-SEP-1997  
 DEFINITION Multidrug resistance relating gene derived from human normal cells.  
 ACCESSION E02326  
 VERSION E02326.1 GI:2170561  
 KEYWORDS JP 1990100680-A/1.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 4378)  
 AUTHORS Ueda,K. and Komano,T..  
 TITLE HUMAN NORMAL CELL-DERIVED MDR RELATED GENE  
 JOURNAL Patent: JP 1990100680-A 1 12-APR-1990.  
 COMMENT  
 OS Homo sapiens  
 PN JP 1990100680-A/1  
 PD 12-APR-1990  
 PE 05-OCT-1988 JP 1988251475  
 PI UEDA KAZUMITSU, KOMANO TORU  
 PC C12N15/12,C12N1/21,C12Q1/68;  
 CC strandedness: Single;  
 CC topology: Linear;  
 CC source: tissue-type=Adrenal gland;  
 CC \*source: clone=SAM1132;

```

FH Key Location/Qualifiers
FH 5'UTR 1..137
FT CDS 138..3980
FT mat_peptide 138..3977
FT /gene='Multidrug resistance relating gene' FT
FT /gene='MDR1' 3981..>4378.
FEATURES
source 1..4378
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 1315 a 818 c 1057 g 1168 t
ORIGIN
Query Match 31.2%; Score 605.4; DB 6; Length 4378;
Best Local Similarity 65.6%; Pred. No. 3.4e-136;
Matches 917; Conservative 0; Mismatches 471; Indels 9; Gaps 2;
QY 162 caggttgaccctgtatattatgttgaaatagtgctgcttgatatttggttaacata 221
Db 473 CAGGTATGCCCTATTATTATACAGTGAATTGTGCTGGGCTGTGCTGTGCTTACATTCA 532
QY 222 gatcttccttgatattataatcagcagcagcagcagcagcagcagcagcagcagcagc 281
Db 533 GATTTCAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592
QY 282 tcaatcagcttggtgcaagacatcgctgctgttgatgctgtgacatcgctgaactaa 341
Db 593 TCATGCTAATATGCCAGACAGAGATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652
QY 342 cactcgatcagacagacattg---acaaatcagtgatgtglatgtgagataagattgctc 398
Db 653 CACCCGACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712
QY 399 gttgttcaaaacatgctactcttttgcttgctgctgctgctgctgctgctgctgctg 458
Db 713 GTTCTTTCAGTCAATGCAACATTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 772
QY 459 gaactcaccatgactgactcactcactcactcactcactcactcactcactcactcactc 518
Db 773 GAAGCTAACCCTGTGTATTTTGGCCATCAGTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 832
QY 519 ttctagatgtgcatcctcactcactcactcactcactcactcactcactcactcactc 578
Db 833 GGCAGAAATACATATCTTCACTTATGATTAAGAACTCTTAAGCGTATGCAAAAGCTGAGC 892
QY 579 tgtgcaagaagaagctgtgctatcactcactcactcactcactcactcactcactcact 638
Db 893 AGTAGCTGAAGAGGCTTGTGGCAGCAATTAAGCTGTGATTTGGATTTGGAGCAAAAGAA 952
QY 639 agaactcacaagatatacacaagaatcacaagaatgaagaagatttgatcataaaagagc 698
Db 953 AGAAGCTGAAGAGTACAAACAAAATTTAGAGAGAGCTTAAGAGATTTGGATTAAGAAAC 1012
QY 699 tatagctcaaaagtgctccttgctgtgctgtgctgtgctgtgctgtgctgtgctgtg 758
Db 1013 TATTACAGCCAAATTTCTATAGAGTCTGCTTCCGTGATGATGATGATGATGATGATG 1072
QY 759 tgccttttgatgatacctccttgatccttcaatgagaagactgataacatcogggac 818
Db 1073 GGCCTTCTGTGATGAGACACCTTGTGCTCTCAAGGGA-----ATATTCTATTGGACA 1126
QY 819 tgtcttgctgttcttcttgaatgaatcactcactcactcactcactcactcactcactc 878
Db 1127 AGTACCACTGTATTTCTTCTGTATTAATTTGGGCTTTTATGTTGGACAGCACTTCC 1186
QY 879 tcaacttgaacactcgaatagccgagagagctgcttcaatlatlccagattatga 938
Db 1187 AAGCATTTGAAGCATTTGCAAAAGCAAGAGAGAGAGCTTATGAATCTTCAAGATTAATGCA 1246

```



DB	Accession	Sequence	Score	DB	Length	Gaps
OY	939	taagaaccocagatagataacttctccacagctgtagatataaactgatacatcatgaagg	998			
Db	1247	TAATTAAGCCCAAGATATTGACACGCTATTTCGAAGAGTGGGCAACAAACGATTAATTAAAGG	1306			
OY	999	aactgtgaatttaanaaagtgtcttccaatatcatcaagaacatcatcaagattct	1058			
Db	1307	AAATTGGGAATTCAGAAATGTTCCACTTTCAGTTACCCATCTCGAAAAAGATTAGATCTT	1366			
OY	1059	gaaagcttgtaactcagaatlaaagtcctgagaagacagtcgccttgctgcgtcaatg	1118			
Db	1367	GAAGGCTCGAAACCGAAGAGTGCAGAGTGGGCGACAGGCTGGCCCTGTTGGAAAACTGG	1426			
OY	1119	cagtgagaagagtaeagtgatgccagctctcagagtgatataatgctccgatagagctt	1178			
Db	1427	CTGTGGGAAGACCAACACAGTCCACCTGATGAGAGGCTCTATGACCCACAGAGGGGAT	1486			
OY	1179	tatcatgtgtgtagaagatgacatcagaagctttaatgtgcgcattatcgagacat	1238			
Db	1487	GGTCAGTGTGATGAGACAGGATATTAGCAACCTTAATGTAGGTTTCTACGGGAATCAT	1546			
OY	1239	tggagtggtltagaagaagagctgttttctggaagaccacatcagtaaacatataagta	1298			
Db	1547	TGGTGTGGTAGTACGAGAACCTGTATTGTGTTCCACACAGATAGCTGAAAAATTCGGTA	1606			
OY	1299	tggagagatagatgtgactgatagaagatgtagagagacgaaggaagacaatgagta	1358			
Db	1607	TGGCCGCTGAAAAATGTCACCATGAGATGAGATTGACAAAGCTGTCAAGGAAGCCCTTA	1666			
OY	1359	tgaattatcatgagagtttccataaatttaatacatgtgtagagggaagaaagagctca	1418			
Db	1667	TGACTTTTATCATGAACACGCTCATTAATTTTGACACCCTGTTGGAGAGAGGGGGCCCA	1726			
OY	1419	aatgagtgagaggcagaagaacagagatgcgcaatgtcgtgcttagttcgaaacccaa	1478			
Db	1727	GTTGAGTGGTGCGCAGAGACAGAGATGCGCATTCACAGTGCCTGTTGCTGAACCCCA	1786			
OY	1479	gattctgattttaagtatgagctgaagctgcgcctcgatcatcagaagcaagtcagctgtca	1538			
Db	1787	GATCCTCTCGCTGGATGAGGCCACGCTCAGCCTTTGGACACAGAAACGAAAGCATGGTTCA	1846			
OY	1539	agctgcactgagaaag 1555				
Db	1847	GGTGGCTCTGGATTAAG 1863				
RESULT	9					
LOCUS	149610	149610 4646 bp DNA				
DEFINITION		Sequence 2 from patent US 5641508.				
ACCESSION	149610					
VERSION	149610.1	GI:2471830				
KEYWORDS						
SOURCE		Unknown.				
ORGANISM		Unknown.				
REFERENCE		1 (bases 1 to 4646)				
AUTHORS		Li, L. and Lishko, V. K.				
TITLE		Method for delivering melanin to hair follicles				
JOURNAL		Patent: US 5641508-A 2 24-JUN-1997;				
FEATURES		Location/Qualifiers				
	source	1..4646				
BASE COUNT	1371 a	892 c	1129 g	1254 t		
ORIGIN						
Query Match	31.1%;	Score 603.8;	DB 6;	Length 4646;		
Best Local Similarity	65.6%;	Pred. No. 8.4e-136;				
Matches 916;	Conservative 0;	Mismatches 472;	Indels 9;	Gaps 2;		
OY	162	caggtgacccctgatatgttggaatagggtgtcgtcctcgatatttggtatcataca 221				
Db	760	CAGGATAGCCATTTATTATACAGTGAATGTGCTGGGGTGGTGGCTTGCCTTAAATCA 819				

OY	222	gattcctcttgaggttaataactcagacagacagaccagagattctgaaacagtttt	281
Db	820	GGTTCACTTTTGGCTGGCTGGCACTGGAMGACAAATACCAAAATATTAGAAACAGTTT	879
OY	282	tcattcagtttttcgacagagacacatcgctggttttgaatgactgtgacatcgtgtgaacttaa	341
Db	880	TCATGCTATATGCGACAGAGATATAGGCTGGTTGATGTCACGATGTTGGGACCTTAA	939
OY	342	cactcgcatagcagacattgac---aaatcagtgatgttatgtgaagttaagtgcct	398
Db	940	CACCCGACTTACATATGATGATCTCTAAGATTAAATGAAATTTATTTGGACAAAAATTGGAT	999
OY	399	gtgtcttcaaaacgtctactcttttcgatattggtccgtgagctgtgtttgtgtgaagctgt	458
Db	1000	GTTCTTTTCAGTCATATGGCAACATTTTTTCACTGGGTTTATATAGATTTTACACGTGTTG	1059
OY	459	gaaactcaacctagtgactctatccacgtctccctctatatagtcttaagcggcagcatg	518
Db	1060	GAAGCTAACCCCTTGATGATTTTGGCCATCACTCTGTTCTTGGACGTCAAGCTGCTGG	1119
OY	519	ttctagaagtgtcatccatctagccagtaaggaattaaigtcttatccaaagctggtg	578
Db	1120	GGCAAGATACTATCTTCATTACTGTATTAAGACCTTTAGCGTATGCAAAAGCTGGAGC	1179
OY	579	tgtggcagaagaagtcctgttcataatccgacagtcataagcctttlagggccagga	638
Db	1180	AGTAGCTGAAGAGGCTTTGGCAGCAATATGAACTGTGATTTGCATTTGGGAGCAAAAGAA	1239
OY	639	agaactcaagagatcatcaagaatctcaaaagtgtgaagagtttggataaanaagac	698
Db	1240	AGAACTTGAAAGGTACACAAATAATTTAGAACAGCTTAAGAAATTTGGATTAAGAAAGC	1299
OY	699	tatagtctcaaaagtgtctctgtgtcgtgtgactctttatgaaatgaaacctatgact	758
Db	1300	TATTACACCAATATTTCTATAGGTGCTGCTTCCGTGATCTATGCACTTTATGCTCT	1359
OY	759	tgctttttgttatgtgaacctcccttgattcttctatgtgaagacctgtatatacatcgtgac	818
Db	1360	GGCCTTGATGTATGGGACCAACTGGGCTCCCTCAGGGG- ----ATATCTATTGGACA	1413
OY	819	tgcttgccttttcttcttctagtgtaatccatagcaatttcgatgtgagccagctccc	878
Db	1414	AGTACTACTGTATTTCTTTCTTGCTGATTAATTTGGGGCTTTAGTGTTGGACAGCACTGCC	1473
OY	879	tcaacttgaaaccttcgcgaatagccgagagagcttgcccttcataatlttccaggttatga	938
Db	1474	AAGCAATTGAAACATTTCCAAATGCAAAAGAGGACAGCTTTGAAATCTTCAAAATATTGGA	1533
OY	939	taagaaacccagtatagataacttttccacagctgtgatataaacctgaaatccatatgaag	998
Db	1534	TAATATAGCCAAGTATATACAGCTATTTGGAAGATGGGGCAAAACCAAGATAATATAAGG	1593
OY	999	aacttgtgaatttaaaatgtttcttccaattatccataagaagacatctatcaagattct	1058
Db	1594	AAATTTGGAATTCAGAAATGTCTCACTCAATGTACCCATGTCCAAAAGAAAGTTAAATCTT	1653
OY	1059	gaaagttcgaatctcgaatlaagtcctgtgaaggaagcgtccgtgtgtgcgtcctaaagg	1118
Db	1654	GAAAGGCTTGAACTTGAAAGGTGCAGATGGGCGACAGGAGTGCGCTGTGGTGGAAACAGTG	1713
OY	1119	cagtgtgaagaagtaacgttagtccaagctcttcgcagaggttatatgataccgatactgact	1178
Db	1714	CTGTGGGAAGGACACACAGTCCAGCTGAGTGAAGAGGCTCTATGACCCCAAGAGGGGAT	1773
OY	1179	tatcatgtgtgatagagaatgacatcagagctttaaatgtgcygcattatcgagaccatat	1238
Db	1774	GGTCAAGTGTATGAGACAGGATATATAGGACCAATAAATGTAAAGTTTCTTACGGGAATCAT	1833
OY	1239	tggagtggtgtgaagaagctgtttgtgttcgagaccacatcaggtlaaanaataataagta	1298
Db	1834	TGTTGTGTGATGAGAACTGTGATTTGTTTGGCCACACGATATGCTGAAACATTTCCGTA	1893



OY	1299	tggagagaatgatgtcgcactgatagaagagatagtgaagagacgaaggaagcaaatgcycta	1358
Dd	1894	TGGCCGCTGAATAATGTTCACCATGATGAGATTGTAGAAGACTGTCAAGAGGCCAATGCCTTA	1953
OY	1359	tgatttcatcattgagggttttcctcaataattcatatcattgttagtgagggaaaagagctta	1418
Dd	1954	TGCACCTTTATCATTGAACAACCTCGCTCATTAATTTTACACCTCGTTGGANAGAGAGGGGCCCA	2013
OY	1419	aatgagtgaggagggcagaacaacagagatcgcaattgctcgtgtcattgatttcgaaccocaa	1478
Dd	2014	GTTTGAGTGGTGGGCGCAGAAAGCAGAGATGCGCCATTGTCACGTCGTCGTTGCCAACCCCAA	2073
OY	1479	gatttcgattttagtgttggtgcactgcgtccctcgtgattccagaagcaagtcagctgttca	1538
Dd	2074	GATTCCTCTCGCTGAGTAGGGCCACCTCAGCCCTTGACACAGAAACCAGAACGATGGTTCA	2133
OY	1539	agctgcactgagagaag 1555                 	
Dd	2134	GGTGCCTCTGATTAAGG 2150	
RESULT	10		
HUMMDR1			
LOCUS	HUMMDR1	4646 bp	mRNA PRI 03-DEC-1999
DEFINITION	Homo sapiens P-glycoprotein (PGY1) mRNA, complete cds.		
VERSION	M14758		
KEYWORDS	M14758.1 GI:187468		
SOURCE	P-glycoprotein; drug resistance protein; transport protein. Human drug-resistant cell line KB-C2.5 cDNA to mRNA, clones lambda-HDR(10, 5, 104).		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 4646) Chen,C.J., Chin,J.E., Ueda,K., Clark,D.P., Pastan,I., Gottesman,M.M. and Roninson,I.B. Internal duplication and homology with bacterial transport proteins in the mdrl (P-glycoprotein) gene from multidrug-resistant human cells		
TITLE	Cell 47 (3), 381-389 (1986)		
JOURNAL	Medline 87028230		
REFERENCE	2 (sites)		
AUTHORS	Ueda,K., Clark,D.P., Chen,C.J., Roninson,I.B., Gottesman,M.M. and Pastan,I.		
TITLE	The human multidrug resistance (mdr1) gene. cDNA cloning and transcription initiation		
JOURNAL	J. Biol. Chem. 262 (2), 505-508 (1987)		
MEDLINE	87109132		
REFERENCE	3 (bases 971 to 985; 3095 to 3109) Kloka,N., Tsubota,U., Kakehi,Y., Komano,T., Gottesman,M.M., Pastan,I. and Ueda,K.		
AUTHORS	P-glycoprotein gene (MDR1) cDNA from human adrenal: normal P-glycoprotein carries Gly185 with an altered pattern of multidrug resistance		
TITLE	Biochem. Biophys. Res. Commun. 162 (1), 224-231 (1989)		
JOURNAL	Medline 89322246		
MEDLINE	[2] sites.		
COMMENT	Draft entry and computer-readable sequence [1] kindly submitted by I.B.Roninson, 13-AUG-1987. The sequence shown is of a cDNA clone initiating at a minor upstream transcription initiation site and containing the major site of transcription initiation. Location/Qualifiers 1..4646 /organism="Homo sapiens" /db_xref="taxon:9606" /map="7q21" 1..4646 /gene="PGY1" 382 /gene="PGY1" /note="g in [1]; a in [2]" 425..4267		
FEATURES	Source		
CDS	variation		

[illegible]

QY	399	gtgttttcaaacaatgtctactcttttcgattgctgctgacgtgtgtttgtgtgaagggctg	458
Db	1000	gtttttttcagtcacatggcacaacattttttcactggtggtttatgtaggattttacacgtggttg	1059
QY	459	gaacatcaccttagtgactctatccacgctcctccttaataatggtcttaagcggacatg	518
Db	1060	gaacctaaccttctgtatttttggccatcagctctgttcttgacactgctgactgctg	1119
QY	519	ttctaagatggtcatctcatctatgacccaaggaatgaatgaatgcttaatccaaagctgggc	578
Db	1120	ggcaaaagatctatcttacttacttactatgataaagaaacctgttagcgatgcaaaagctggagc	1179
QY	579	tgtagcagaagaagctctgtcatcgaatccgaacagtcataagcctttagggcccaagaa	638
Db	1180	agtagctgaagagctcttgccacacatttgaactgtgattgtcattttggagacaaaagaa	1239
QY	639	agaactccaaggtatcacacagaatctcaaaagatgaagatcttggcataaaagagc	698
Db	1240	agaaacttgaaagggtacacaaaattttagbagaagctaaagaaattgggattaaagaaagc	1299
QY	699	tatagtctcaaaagtgtctctgtgtgctgtgtaactctttaaagaaatggaaactatggact	758
Db	1300	tatttacagccaattttcttattagtgctgctgttccgcgatcatttgatcttattggcct	1359
QY	759	tgctttttgtatggaacctcccttgatcttctaagagaacctgtatataccatcggagc	818
Db	1360	ggccttctggtatgggacccactttggcttccctcagggga-----atattttattggaca	1413
QY	819	tgcttgcgtcttttcttcttagtgtaatcagaagctatgcatcttggaagcaagctcc	878
Db	1414	agtagctactgtattttcttttctgtattatttgggcctttttagtggtagacagcatctcc	1473
QY	879	tcacttgaacacctgcgaatgcccgaagagctgctcttcatatcttccaggtatatga	938
Db	1474	aagcattgaaacattttgcaaatgcagaagagagcagcttatgaaattctcaagatattga	1533
QY	939	taagaacccaagtatagataactcttccacagctgtgatatatacctgataatccatagaag	998
Db	1534	taattaaagccaagtattttacacgctattttcgaaagatggggcacaacacagataattatgaag	1593
QY	999	aacgttgaatttaaaaatgtcttcttcaatataccataagaccatatacaagattct	1058
Db	1594	aaatttggaaatttcgaaattgtttactttcagatttaccatttgcaaaaagaaatttaagatttt	1653
QY	1059	gaaggtctgaatctcgaatctaagcttaagctgagagaagcgtgctgtgcttccaaatg	1118
Db	1654	gaagggcctgaacctgaaggtgcagagatggggcagacggtggccctgttggaacaagatgg	1713
QY	1119	cagtgggaagaagtagcgttagtccagctctcgaagagtataatgctcsgatgatgctt	1178
Db	1714	cttggtggaaagcacaacactccagctgattgacagagccttattgacccccacagaggggat	1773
QY	1179	tataatgigtatgataagataacatcagagcttaaatggcgactatcgagaccat	1238
Db	1774	ggtcagcttgcattgacacagatattatggacacataaattgtaagctttttcagcggaaattcat	1833
QY	1239	tggaagtgttaagtcacaagcctgtttgttgcggaccacacacgaataacaatatcaagta	1298
Db	1834	tggtgtgtgtgacaggaacctgtattgtttggcacacacgatagctgaaaaacatttcgcta	1893
QY	1299	tggaagagatgatagtcactgataagaagatggaagacgaagaaggaanaatgctga	1358
Db	1894	tggcctcgtaaaatgtcacacatgcatgataagatttgaaagctgtccaagaaagccaatgacct	1953
QY	1359	tgatttatacatggagtttccataaatttaatacatatggtgaagggaaaaaagagctca	1418
Db	1954	tgactttattcttgaaactgctctattaaattttgacacccctgtttgagagagagggggccca	2013
QY	1419	aatgatagtgaaggaagaacaagagatcgaaattgtctgtgcttaagtcgaaaacccaa	1478
Db	2014	gttagagtgctggcgacagagacagagatccgcatttgacagctggccttggttccaaacccaa	2073
QY	1479	gattctgattttagatgaggtctagctctgcctcctggtatlaagaagaagatcagctgttca	1538

Db	2074	GAFTCTCTGCTGGATGAGGCCACGCTTGGACACAGAAAGCAAGCAGTGTCTCA	2133
QY	1539	agctgcactgcagaag 1555	
Db	2134	GGGGGCTGTGATPAAAG 2150	
RESULT 11			
AR091275	AR091275	4669 bp	DNA
LOCUS	Sequence 1	from patent US 5994088.	PAT
ACCESSION	AR091275		07-SEP-2000
VERSION	AR091275.1	GI:10018030	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 4669)		
AUTHORS	Mechetner, E. and Roninson, I.B.		
TITLE	Methods and reagents for preparing and using immunological agents specific for P-glycoprotein		
JOURNAL	Patent: US 5994088-A 1 30-NOV-1999;		
FEATURES	Location/Qualifiers		
source	1..4669		
BASE COUNT	1393 a 894 c 1130 g 1252 t		
ORIGIN			
Query Match	31.1%;	Score 603.8;	DB 6;
Best Local Similarity	65.6%;	Pred. No. 8,4e-136;	Length 4669;
Matches 916;	Conservative 0;	Mismatches 472;	Indels 9;
			Gaps 2.
QY	162	caagttgaccctgatatgttggaatagtggtgcgcctgatttgyttacata	221
Db	760	CAGGTATGCCCTATTATTAAAGTGAATTGTCTGGGGGCTGGTTCCTTACATTCA	819
QY	222	gatttcctgtgataataactgcagcacgacagacaaagagaltcgaaaaagtltt	281
Db	820	GGTTTCATTTTGGTGGCTGGCAGCTGGAACAAATACAAATAGAAAACAGTTT	879
QY	282	tcattcagttttgycacagagacatcgycgtgttgatagctgtgacatcggtgaactaa	341
Db	880	TCATGCTATTATGCGCAGAGATAGCGTGGTTGATGTCACGAGATGTTGGGAGACTTAA	939
QY	342	cactgcacatgcagaagacatgtt---acaaaatcagtaagttgatgaagaatgaagctct	398
Db	940	CACCCGACTTACAGATGATGCTCCACAGATTAATGAGAAATTTGGTGCACAAATTTGGAAT	999
QY	399	gttgtttcaaaaacatgtctcacttttttcgaattgycgcgtgagttggtttgtaaggctc	458
Db	1000	GTTCTTTACGTCAATATGCAACAATTTTTCACCTGGGTTTATAGTAGAATTACACGTGGTTG	1059
QY	459	gaaactcacocctagtgactctatccacgctctccctctataatgactcagcgccagcgacatg	518
Db	1060	GAAAGTAAACCTTTGTGATTTTGGCCATCAGTCTCTGTTTGGACGCTCAGCTGTCTGTCG	1119
QY	519	ttctagatagtgatcatctcctctgacccaagtaagaatgaatgaatgctctccaaagctgggac	578
Db	1120	GGGAAAGATACATCTTCTTCAATTTACATGATTAAGAAACCTTTAGGCTATGCAAAAGCTGGAC	1179
QY	579	tgtgacagaagaagctctgtcatcaatccgaacagctacagctcttaagagccacagaagaa	638
Db	1180	AGTAGCTGAAGAGCTTTGGCAGCAATTAGAACTGTGATTTGGAGGACAAAGAAAGAA	1239
QY	639	agaacttcaaaagtlacacagaatctcaaaatgcagaagatlttggcataaaaagac	698
Db	1240	AGAACTTGAAGAGTACAAACAAATTTTAAAGAAAGCTTAAACAAATTTGGATPAAAGAAAC	1299
QY	699	tatagcttcaaaagtctcctgtggtcgtgtgtgactctttagaatggaaactatagact	758
Db	1300	TATTATAGGCCAATATTCTTAAGTGTCTCTTCCGCGATGATTAAGCAATCTTATGCTGTCT	1359

759 tgccttttgatggaaccctctatcttgaatgagaacctgatataccatcgagac 818  
1360 ggccttcttgatggaaccctcttgcctctcagaggga-----ATATTCTATTGGACA 1413  
819 tgtcttgctgttcttctttagtgaatccatagcagttatgcatgtgagcagctccc 878  
1414 AGTACTCACTGATTTCTTTCTGTATTAATGGGGCTTTTACTGTTGGACAGCATCTCC 1473  
879 tcaacttgaaccttcgcaatagcccgagagagctgccttcaatlttccaggttatgta 938  
1474 AAGCATTGAAGCATTTGCAAAATGCAAGAGGACACCTTATGAAATCTTCAATATTGA 1533  
939 taagaaccacagatagaacttaacttccacagctgatatgaataactgatatgaag 998  
1534 TAATTAAGCCACTATTGACAGCTATTTCAGAGTGGGCAACAAACAGATAATTAAAGG 1593  
999 aacttggaatttaaaatgttcttcaatatacatcaataacatcatcaaatctc 1058  
1594 AATTTGGAAATTCAGAAATGTTCACTTCAGTTACCATCTGAAAAAGATTAAAGTCTT 1653  
1059 gaaaggtcgaatccagaataatagctcggagagacagtcgccttggtcgtcgaatg 1118  
1654 GAAGGCGCTGAACCTGAAGGTGACAGAGTGGCAGACGGTGGCCCTGTTGGAAAGAGTGG 1713  
1119 cagtggaagaagtaagcgtagctcagctctcggagaggtatatagtatcgatgagctt 1178  
1714 CTGTGGAGAGACACACACAGTCCAGCTGATGACAGAGGCTCTATGACCCACAGAGGGGAT 1773  
1179 tatcatggtgagtgagaaatgacatcagagctttaaattgtcggcatltaacagacat 1238  
1774 GGTCAAGTGTGATGACAGGATTAAGACCATTAATGTAAGTTCATGAGGGAATCAT 1833  
1239 tggagtggttagtcaagagcctgcttcttggtcgagaccacatcagtaacatataagta 1298  
1834 TGGTGTGGATCAGGAACCTGTATTGTTGCCACACAGATAGCTGAAAAACATTGCGCTA 1893  
1299 tggagagatagtgatcagatgagaagatgagaagagagaaggaaggaatgagcgt 1358  
1894 TGGCGGTGAATGTCACCATGATGAGATGAGAAAGCTGCAAGGAAGCAAAATGCTTA 1953  
1359 tgatttcatcagtaggttcttctaataatlaatacatatgtaggaggaagagagctca 1418  
1954 TGACTTTATCATGAAACGCTCATTAATTTGACACCCGTTGGAGAGAGAGGGGCCCA 2013  
1419 aatggtggaaggaagaagaacagagatgcgaatctgcgtccttagtgcgaacccaa 1478  
2014 GTTGTGAGTGGGCGAGAGCAGAGATGCGCATTTGACAGTGGCCCTGTTCCGCAACCCCA 2073  
1479 gattctgatttgaatgagagctagctcgcctgagattcagaagaagaagtcagctgtta 1538  
2074 GATCTCTCTGCTGATGAGGCGACGTCAGCTTGGACACAGAAACGAGAGTGGTTCA 2133  
819 tgtcttgctgttcttctttagtgaatccatagcagttatgcatgtgagcagctccc 878  
1414 AGTACTCACTGATTTCTTTCTGTATTAATGGGGCTTTTACTGTTGGACAGCATCTCC 1473  
879 tcaacttgaaccttcgcaatagcccgagagagctgccttcaatlttccaggttatgta 938  
1474 AAGCATTGAAGCATTTGCAAAATGCAAGAGGACACCTTATGAAATCTTCAATATTGA 1533  
939 taagaaccacagatagaacttaacttccacagctgatatgaataactgatatgaag 998  
1534 TAATTAAGCCACTATTGACAGCTATTTCAGAGTGGGCAACAAACAGATAATTAAAGG 1593  
999 aacttggaatttaaaatgttcttcaatatacatcaataacatcatcaaatctc 1058  
1594 AATTTGGAAATTCAGAAATGTTCACTTCAGTTACCATCTGAAAAAGATTAAAGTCTT 1653  
1059 gaaaggtcgaatccagaataatagctcggagagacagtcgccttggtcgtcgaatg 1118

RESULT 12  
LOCUS 108557 4669 bp  
DEFINITION Sequence 3 from Patent WO 8705943. PAT 02-DEC-1994  
ACCESSION 108557  
VERSION 108557.1 GI:588735  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4669)  
AUTHORS Robinson J.B., Pastan, I.H. and Gottesman, M.M.  
TITLE COMPOSITIONS AND METHODS FOR CLONES CONTAINING DNA SEQUENCES  
ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS  
JOURNAL Patent: WO 8705943-A 3 08-OCT-1987;

FEATURES  
Location/Qualifiers  
source 1..4669 /organism="unknown"  
BASE COUNT 1394 a 892 c 1129 g 1254 t  
ORIGIN  
Query Match 31.1%; Score 603.8; DB 6; Length 4669;  
Best Local Similarity 65.6%; Pred. No. 8.4e-136;  
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;  
162 cagcttgacacggtatattatgtggaatgagtgctgctgcatltagtgaatataca 221  
760 CAGGATAGCGCATTTTATTAAGTGAATGCTGGGGTGGCTGCTTCACTTAACTTCA 819  
222 gattccttgatgataatactgacgacagacagacccaagagattcgaaacagtttt 281  
820 GGTTCATTTTGTGCTCCGACGCTGGAAGACAAATACCAAAATTTAGAAACAGTTT 879  
282 tcaatcagtttgacacagacatcgctggtttagatagctgtgacatcggtgaattaa 341  
880 TCATGCTATAATGCGACAGAGATAGGCTGTTGATGTCACAGATGTTGGGAGCTTAA 939  
342 cactcgcatgacagacatgac---aaatcagtgatggtatggaatgaatgctct 398  
940 CACCCGACTTACAGATGATGATGCTCTAAGATTAATGAAGTTATGGTGACAAATTTGGAAT 999  
399 gtgttcaaaaacatgctacttcttgcagatgagctgagctggttggtagaaggctg 458  
1000 GTTCTTTCACTGATGAGCAACATTTTTCACCTGGGTTTATAGATGATTAACAGTGGTGG 1059  
459 gaaactccacagtagactatccacgctctctctatgaatgagcttgaagcgagacatg 518  
1060 GAAGCTAACCCCTTGATTTGCGCATACGCTGCTTGGACACTGTCAGCTGCTCTG 1119  
519 tctcagtagtgcacatcattgacaggaatgaatgaatgcttccaaacgctgggac 578  
1120 GCCAAAGATACATCTTCTTACATGATTAAGACACTTACGGTATGCAAAAGCTGGAC 1179  
579 tgtgcagaagaagctctgtgcatcaatccgaacagtcacatagccttagggccagagaa 638  
1180 ACTACTGTAAGAGGCTTGGCAGCAATTAAGACTGTGATTCGATTGGAGGCAAAAGAC 1239  
639 agaacttcaaaagtataccacagaatctcaagatgcaagagtttggatataaaagac 698  
1240 ACAACTGTAAGAGTACAAACAAATTTAGAAAGACTTAAAGAAATTTGGATTAAGAAC 1299  
699 tatagctcaaaagtctctgtgctgtgctgtgacttcttcaatgaaatgaaatgagct 758  
1300 TATTACAGCAATTTTCTATAGGTCCTCTTCCGCTGATCTATGATCTTATGCTCT 1359  
759 tgccttttgatggaacctctgattcttaatgagagaacctgatalacatcaggagac 818  
1360 GGCCTTCTGTGATGGGACACCTTGCTCTCAGGGGA-----ATATTCTATTGGACA 1413  
819 tgtcttgctgttcttctttagtgaatccatagcagttatgcatgtgagcagctccc 878  
1414 AGTACTCACTGATTTCTTTCTGTATTAATGGGGCTTTTACTGTTGGACAGCATCTCC 1473  
879 tcaacttgaaccttcgcaatagcccgagagagctgccttcaatlttccaggttatgta 938  
1474 AAGCATTGAAGCATTTGCAAAATGCAAGAGGACAGCTTATGAAATCTTCAATATTGA 1533  
939 taagaaccacagatagaatcttccacagctgataataaactgaaatccatgaag 998  
1534 TAATTAAGCCAAATGTAAGAGCTATTGCAAGAGTGGGCAACAAACAGATAATTAAAGG 1593  
999 aacttggaatttaaaatgttcttcaatatacatcaataacatcatcaaatctc 1058  
1594 AATTTGGAAATTCAGAAATGTTCACTTCAGTTACCATCTGAAAAAGATTAAAGTCTT 1653  
1059 gaaaggtcgaatccagaataatagctcggagagacagtcgccttggtcgtcgaatg 1118

Db 1654 GAAGGGCTGAACCTGAGGTGACAGTGGGACAGCGGTGGCCCTGTTGGAAACAGTGG 1713  
QY 1119 cagtgaggaaagatagcagtgatccagcttcctgcagaggtatataatcgcagatgagctt 1178  
Db 1714 CTGCGGGAAGGACACACAGTCCAGCTGATGACAGAGGCTCTATACCCACAGAGGGGAT 1773  
QY 1179 tatcatgtgtatgataagatcagatcagctttaaattgtgcgcatatcagagcatat 1238  
Db 1774 GGTGAGTGTGATGACAGAGATATTAGACCATTAATGTAAGTTTGTTGACGGAAATCAT 1833  
QY 1239 tggagtggtatgataagagcctgtttgttcgagagcaacccatcagtaataatcagta 1298  
Db 1834 TGGTGTGTGTGATGAGGACCTGATTGTTGGCCACGATGACCTGAAACATTCGCTA 1893  
QY 1299 tggagcagatgatgtgactgtatgataagagatgagagcaagaaggagaaatgtgta 1358  
Db 1894 TGGCGGTGAATAATGTCACATGATGATGATGAGAAAGCTTCAAGAGAACCATGCTTA 1953  
QY 1359 tgatttatcatgtgagttccttaataatataatcatalgtgtaggggaaaaaggagctca 1418  
Db 1954 TGACTTATCATGAAACGCTCATTAATTTGACACCCCTGTTGAGAGAGAGGGGCCCA 2013  
QY 1419 aatgagtgaggagagcaagaacagagatcgcacatgttcgttcgttgcgttgcgttgc 1478  
Db 2014 GTTGAAGTGTGGGAGAGAGAGATCGCCATTCGACAGTCCCTGTTCCAAACCCCA 2073  
QY 1479 gatctgatttatgataagagctatgctgtcctgattcagaaagcaagatgagctgtca 1538  
Db 2074 GATCTCTCTGCTGATGATGAGGCGACGCTCAGCTTGGACACAGAAAGCAGAGAGTGTCA 2133  
QY 1539 agctgcaactgagaagg 1555  
Db 2134 GGTGGCTCTGATGATAGG 2150

RESULT 13  
AR028671  
LOCUS AR028671 6505 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 5 from patent US 5858744.  
ACCESSION AR028671  
VERSION AR028671.1 GI:5940644  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 6505)  
AUTHORS Baum, C., Stocking-Harders, C. and Osterag, W.  
TITLE Retroviral vector hybrids and the use thereof for gene transfer  
JOURNAL Patent: US 5858744-A 5 12-JAN-1999;  
FEATURES  
source 1..6505  
BASE COUNT 1719 a 1506 c 1606 g 1674 t  
ORIGIN

Query Match 31.1%; Score 603.8; DB 6; Length 6505;  
Best Local Similarity 65.6%; Pred. No. 8.5e-136;  
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;

Db 2332 CACCCGACTTACAGATGATGCTCTAAGATTAAGATTATGTTGATGACAAATTTGAAAT 2391  
QY 399 gttgttcaaaacatgctacttttttgatgagcctgaggttggttggaaaggctg 458  
Db 2392 GTTCTTTGATGCAATGGCAACATTTTTCACGTTTATATATAGATTTTACACGTGGTTG 2451  
QY 459 gaaactcaccttagtgcactcatalccacgctcctccttataatggttcacgcgagcatg 518  
Db 2452 GAACCTAACCCCTTGATGATTTTGGCAATCAGTCGTTCTTGATGATGATGATGATGATG 2511  
QY 519 ttctagatgtgtcatcattcattgaccagtaagaaattagtgctattccaaagctgggac 578  
Db 2512 GGCAAAGTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 2571  
QY 579 tggcagaagaagctgtgtcatcattccagagagctatagcctttagggccagagaa 638  
Db 2572 AGTAGCTGAAGAGGTCTTGGCAGCAATTAGAAGCTGTGATTTGAGAGACAAAGAA 2631  
QY 639 agaacttcaagagatatacacagaaatcccaagatgcaaaagatttggcataaaaggac 698  
Db 2632 AGAAGCTGAAGGATNCAACAAATAATTTAGAAAGATTAAGAAATTTGGAGTAAGAAAG 2691  
QY 699 tatagctcaaaagtgctccttggtgtgctgtgcttcttattgaaatgaaactatgact 758  
Db 2692 TATTACAGCCAAATTTCTATAGGTGCTGCTTCTGCTGATCATCTATCTATCTCT 2751  
QY 759 tgccttttgatgaaacccctctgattccttaatgaggaacctgatatccacggagac 818  
Db 2752 GGCCTTCTGATGAGGACACCTTGCTCTCTCAAGGGA-----ATATTCTATTGGACA 2805  
QY 819 tgtcttgctgttcttctttagtgaatccatagagattatgcatlgyagagagctccc 878  
Db 2806 AGTACTGACTGATTTCTTTCTGATTAATTTGGGCTTTTATGTTGAGAGCATGCTCC 2865  
QY 879 tcacttgaaccccttgcaatagccagagagctgtccttccatatttccaggtattga 938  
Db 2866 AAGCATTGAAACATTTGCAATTTGCAAGAGGAGGAGGAGCTTATGCAATCTCAAGATTTGA 2925  
QY 939 taagaaaccagatgataatccttccacagcctggatataactgaaatccatcaagag 998  
Db 2926 TAAATAGCCAAATTTGACACTTATTCAGAGATGAGGACAAACAGATTAATTTAAGGG 2985  
QY 999 aactgtgaatttaaaatgttcttccatataatccaaagacatcatcaaatctt 1058  
Db 2986 AAATTTGGAATTTCAAGAAATGTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 3045  
QY 1059 gaaagtgctgataatcagaaatgaatgctggagagacagtgctgtgctgtcctcaatg 1118  
Db 3046 GAAGGGCTGAAACCTGAAGGTGACAGAGTGGGCAACAGGTGGCCCTGTTGGAAACAGTGG 3105  
QY 1119 cagtgaggaaagatagcagtgatccagcttcctgcagaggttataatcgcagatgagctt 1178  
Db 3106 CTGTGGGAGAGACAAACAGTCCAGCTGATGATGATGATGATGATGATGATGATGATG 3165  
QY 1179 tatcatgtgtatgataagatcagatcagctttaaattgtgcgcatatcagagcatat 1238  
Db 3166 GGTGAGTGTGATGACAGAGATATTAGACCATTAATGTAAGTTTGTTGACGGAAATCAT 3225  
QY 1239 tggagtggtatgataagagcctgtttgttcgagagcaacccatcagtaataatcagta 1298  
Db 3226 TGGTGTGTGTGATGAGGACCTGATTGTTGGCCACACAGATGATGATGATGATGATG 3285  
QY 1299 tggagcagatgatgtgactgtatgataagatgagagagcagcaaggagaaatgtgta 1358  
Db 3286 TGGCGGTGAATAATGTCACATGATGATGATGATGATGATGATGATGATGATGATG 3345  
QY 1359 tgatttatcatgtgagttccttaataatataatcatalgtgtaggggaaaaaggagctca 1418  
Db 3346 TGACTTATCATGAAAGCTGCTCATTAATTTGACAACTGCTTGGAGAGAGAGGGGCCCA 3405  
QY 1419 aatgagtgaggagcaagaacagagatcgcacatgtgtcgttgcgttgcgttgcgttgc 1478

QY	DB	1479	gattctgatttagatbaggactgctcgtccctgattacaagaagatcagctgttca	1538
DB	3466	gattccctccctgctgagwagagccacgctgacgctttggacacagaaagagacagtggttca	3525	
QY	1539	agctgcactcggaggaag	1555	
DB	3526	GGTGGCTCTGTGATTAAGG	3542	
RESULT 14				
LOCUS	AX012321	8630 bp	DNA	PAT 06-SEP-2000
DEFINITION	Sequence 2 from Patent EP0955374.			
ACCESSION	AX012321			
VERSION	AX012321.1	GI:99898370		
KEYWORDS				
SOURCE				
ORGANISM	synthetic construct.			
REFERENCE	artificial sequence.			
AUTHORS	1 (bases 1 to 8630)			
TITLE	Baum,C.D., Hildinger,M. and Ostertag,W.P.			
JOURNAL	Retroviral vectors for gene transfer			
	Patent: EP 0955374-A 2 10-NOV-1999;			
	HEINRICH PETTE INST (DE)			
FEATURES				
source	1..8630	Location/Qualifiers		
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="Provirale Plasmid-DNA"			
	1..160			
misc_feature	/note="Plasmid-R ckgat (pUC)"			
	1..8630			
misc_feature	/note="Retroviral-er Expressionsvektor SPbeta91MSA1"			
	161..677			
	/note="5'-LTR"			
	532..1219			
5'UTR	1220..5062			
mat_peptide	/note="msal mdr1 cDNA"			
	5215..5774			
misc_feature	/note="3'-LTR"			
	5775..8630			
misc_feature	/note="Plasmid-R ckgat (pUC)"			
	BASE COUNT 2234 a 2011 c 2189 g 2196 t			
ORIGIN				
Query Match	31.1%	Score 603.8;	DB 6;	Length 8630;
Best Local Similarity	65.6%;	Pred. No. 8,6e-136;		
Matches 916;	Conservative 0;	Mismatches 472;	Indels 9;	Gaps 2;
QY	152	caggttgccccgtgtattgttggaatagtggttctgctcgttattttggttatacata	221	
DB	1555	CAGGTATGCCATATATATACAGTGGATTGGCTGGGGCTGGTGTGCTTACATTCA	1614	
QY	222	gattcccttgatataactaacctgcagcagcagacagcaagagagattcgaacacagtttt	281	
DB	1615	GGTTTCATTTTTGGTCCCTGGCAGCTCGGGAACAAATATACAAATAATTAGAAAAACAGTTT	1674	
QY	282	tcaatcagtlttgycacagacatcgctggttggatagctgttgacatcgtgtaacttaa	341	
DB	1675	TCATCTCTATTAATGACAGAGATAGCGCTGGTTGATGTCACAGATGTTGGGAGCTTAA	1734	
QY	342	cactgcagatgcagacatgac---aaatcgtgatgtgattggaagataagattgctct	398	
DB	1735	CACCCGACTTACAGATGATGCTCTTAAGATTATATGAAGTTATTTGGTGACAAATTTGAAT	1794	
QY	399	gttgcttcaaaacagctactacttttgcagtgcgtcgaggttggttttgtaagagctc	458	
DB	1795	GTTCTTCAGTCAAGTAAAGGCAACATTTTTCACGTGGTTTATAGAGATTATACACGTGGTTG	1854	
QY	459	gaactacacctagtgactctataccaagtlctccttataatgagcttcagcgagacatg	518	

Db	1855	GAAGCTAACCCCTTGATTTGGCCATACAGTCCCTGTTCTTGGACGTGCAGCTGCTCG	1914
QY	519	ttctcgaagtggtcatcctatctatgcacagtaaggaattaaigtcctatctccaagcttgggc	578
Db	1915	GGCAAAAGTACTATCTTCATTCTACTGTATTAAGAAGCTTTAGCCTATGCAAAAAGCTGAGC	1977
QY	579	tgtgcagaagaagtccttgcatactcgaacagtcatacagtcctttaaggcccgagaa	638
Db	1975	AGTACTGAAGAGGCTGTGGCAGCAATTAGAACCTGTGATTCATTTGGAGCAAAAAGCA	2034
QY	639	agaacttcaaaagtcatacagaagatcccaagaatgcagaagatlttggcataaaagac	698
Db	2035	AGAACTTAAAGGATACAAACAAAAATTTAGAGAGAGCTTAAAGAAATTTGGATTAAGAAAC	2094
QY	699	tatagcttcaaaagtcgtctctgtgtcgtgtactcttattgaatgaaagcaatgcagct	758
Db	2095	TATTAACACCAATTTTCTAATAGTGGCTGCTTCTCCTCATCTATAGCATCTTATCTCT	2155
QY	759	tgccttttgatagaaacctctgattccttaatbgaaacctgatalatacaaccgagac	818
Db	2155	GGCCCTTGTGATGAGACACACTTGGTCHCTCAGGGGA-----ATATTTATTTAGACA	2200
QY	819	tgtcttcctgttcttcttagtgttaatccatacagttatgtcatltggagcagcagctcc	878
Db	2209	AGTACTCATCTATCTTCTTCTTCTGTATTAATATGGGGCTTTTAAAGTTAGTTGAGACAGCATCTCC	2266
QY	879	tcaacttgaaccttcgcaatagccgagagctgaccttcatctatcttccaagtattga	938
Db	2269	AAGCTTTAAGCATTTTGCCAAATGCAGAGAGACAGCTTATGAATCTTTCAAGATATTTCA	2322
QY	939	taagaagaccagtatagataactttlccaaagcttgcataataaactgcatacagaaag	998
Db	2329	TAAATAGCCAGATATATGACACTATTCGAGAGAGTGGGCAAAACAGATATATTAAGG	2388
QY	999	aactgtggaatttaaaatgttcttccaattatcgaatcaagacatctataagattct	1055
Db	2389	AAATTTGGAAATTCAGAAATGTTCACTTCACTTACCTTCCATCTCGAAAAAGATTAAAGATCTT	2445
QY	1059	gaagagtcgtgaatctcgaatcaatagcttctgagagagagctgccttgtagtcgtctaaag	1114
Db	2449	GAAAGGCTTGAACCTGAAAGTGTGCAAGTGTGCAAGATGGGCAAGCGTGGCCCTGTGGAAACAGTGG	2500
QY	1119	cagtcggaagagtaacggttagtcacagctcttcgcagaggttatalatgcagatgatgctt	1176
Db	2509	CTGTGGGAAGAGCAACAACGTCCAGCTGATGACAGAGGCTCTATGACCCACAGAGGGGAT	2566
QY	1179	tatcatgtgtgagtcgaagatgcatacagagctttaaattgtgcgcatlatacgaacatat	1233
Db	2569	GCTCAGTCTTATGATGACAGCGATTTATGAGCAACATTAATGTAAAGTTTCTTCAAGGGAAATCAT	2622
QY	1239	tggagtggttgtaagaagcgtgtgtgttcgcggagacacatcagaataaataatacaaga	1296
Db	2629	TGTTGTGGTGTAGTAGGAACCTGTATTTGTTTGGCCACCAGATATGCTGAAGAAACATTCGTCTA	2686
QY	1299	tggagagatagtcgtgactgacgtgagagatcgtgagagcagcaagggaagcaaatgcgt	1356
Db	2689	TGGCCGTGAATAATGTCAACATGATGATGATTTGAGAAAGCTGTCAAGAAAGCCAAATGCCA	2744
QY	1359	tgaattatcatagatcttccataaataattaaatacattgltgtagygggaaagagacta	1416
Db	2749	TGACTTTATCATGAAGCTGCCTCTAATTAATTTGACACCCTGTGTGGAGAGAGAGGGGCCCA	2800
QY	1419	aatgagtgagagagcaagaacagagagatcgcgaattgtcgtgtcgttgaattccgaacccaa	1477
Db	2809	GTTAGTGTGTGGGAGAGAGCAGAGATGCGCATTTGACACGTGGCCCTGTTGGCAACCCCA	2866
QY	1479	gattcgtattatagatagagctacgtctgcctccggaattcagaagaagtaagctgttca	1536
Db	2869	GATCTCTCTGTTGATAGGCGCAAGCTGAGCTTGGACACAGAAAGGAAGCAGTGGTTCA	2922
QY	1539	agctgcactgagaag 1555	

```

Db      2929 GGTGGCTGTGGATTAAG 2945

RESULT 15
AR028672
LOCUS   AR028672 9318 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 6 from patent US 5858744.
ACCESSION AR028672
VERSION  AR028672.1 GI:5940645
KEYWORDS
SOURCE  Unknown.
        Unclassified.
REFERENCE 1 (bases 1 to 9318)
AUTHORS  Baum,C., Stocking-Harders,C. and Osterlag,W.
TITLE    Retrieval vector hybrids and the use thereof for gene transfer
JOURNAL  Patent: US 5858744-A 6 12-JAN-1999;
FEATURES
        Location/Qualifiers
            source
                1. 9318
                    /organism="unknown"
BASE COUNT 2377 a 2217 c 2364 g 2358 t 2 others
ORIGIN
Query Match 31.1%; Score 603.8; DB 6; Length 9318;
Best Local Similarity 65.6%; Pred. No. 8.6e-136;
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;

QY 162 caggttgaacctgtattatgttgtaatagtggtgctgacctgttttggtaacata 221
Db 2111 CAGGTATGCCATTATTAACAGTGAATTTGGTGGGTGCTGCTGCTTACATTCA 2170
QY 222 gattcccttggtgatttaactgacagacagacacgaagagatttcgaacagtttt 281
Db 2171 GGTTCATTTTGGTGCTGGAGCTGGAGCAATACACAAAATTGAAAAAGTTT 2230
QY 282 tcaatcagtttggcagaagacatcgcgtggttgatagctgtgacatcggtgaactaa 341
Db 2231 TCATGCTATATATGCGACAGAGATAGCTGCTTGATGTCACAGATGTTGGGAGCTTAA 2290
QY 342 caccgcgatgcacacattgac---aaatcagtgatgatttgagataaagtgcct 398
Db 2291 CACCGCACTTACAGATGATGCTCTAAGATTAAATGAAGTTATGTTGGTGAACAAATTGGAAT 2350
QY 399 gttttccaacacatgctacttttcgattgacctgacgttggttggtgaaaggtctg 458
Db 2351 GTCTTTTACAGTATGCGACAAATTTTTCACGTGGTTTATAGATTATACAGTGGT 2410
QY 459 gaaactcaacctagtgactatccacgctcctcttataatgagcttcagcggaacatg 518
Db 2411 GAAGCTTAACCTTGTGATTTTGGCCATCAGTCTGTTCTTGAGACTGTCAGCTGCTGTG 2470
QY 519 tttaagatggtcctatcattgacacgaagaaagaaatgaagtcctattccaagctggggc 578
Db 2471 GGCAAAAGTACTATCTTATTTACTGATTAAGAACTCTTACGTAAGAAAAGCTGGAGGC 2530
QY 579 tgtgcaagaagaagctctgtcatcacaatccgaacagtcataagccttlaagggcccaagaa 638
Db 2531 AGTAGCTGAAGAAGGTCTTGGACGATTAAGAACTGATGATTTGGAGGCAAAAAGAA 2590
QY 639 agaactcaaaagatatacagaagaatccaagaatgcaagatttggcataaaaaagagc 698
Db 2591 AGAAGCTGAAGAAGGTACAAAATTTGAAAGAGCTTAAAGAAATTCGATTAAGAAAGC 2650
QY 699 tatagcttcaaaagtgtctctgtgtgctgtgactcttataatgaaatgaacatagact 758
Db 2651 TATTACAGCCAAATATTTCTATAGGCTGCTTTCCTGCTGATCTATGCACTTATGCTCT 2710
QY 759 tgccttttggatgaacctcctgattcttaatgaaacotgataacatcagggac 818
Db 2711 GGCCTTCTGGTATGGGACCACTTGCTCTCAAGGGA-----ATATTCTATTGGACA 2764
QY 819 tgtcttgcgttcttctttagtgaatccatagcagttatgacattgagcagcagctccc 878

```

```

Db 2765 AGTACTACTGTATTTCTTTCTGTATTAATTTGGGGCTTTTAAAGTGTGACAGCATCTCC 2824
QY 879 tcaattgaaaccttcgaatagcccgaggagctgtcccttcataatlltccaagttatga 938
Db 2825 AAGCATTTGAAGCATTTTGCATTAATGCAAGGAGGACGAGCTTATGAAATCTTCAAGATTAATGA 2884
QY 939 taagaaacccaglatagataactlltccacagcttgatataaactgaatccataagaag 998
Db 2885 TAATAAGCCAAAGTATTCACAGCTATTCGAAAGAGTGGGACAAACCAATATTAATTAAGGG 2944
QY 999 aactgtgaatttaaaatglttcttcaatataccatacgaacatcatcaatgattct 1058
Db 2945 AAATTGGAAATTCAGAAATGTTCACTTCACTTACCACTTCGAAAGAAAGTTAAGATCTT 3004
QY 1059 gaaagctcgaatcctcgaatlaagctcggagagacagctgccttggctggtcctaaatgg 1118
Db 3005 GAAAGGCTGAACTGAAAGTGCAGAGTGGGACACAGCTGCTGCTGTTGGAACAGTGG 3064
QY 1119 cagtggaaagagtaagtgatgcagctctcgcagaggttatatagatccgagatgctt 1178
Db 3065 CTGTGGGAAGACACAACAGTCCAGCTGATGACAGGCTCTATGACCCACAGAGGGGAT 3124
QY 1179 tatatggtgatatgaatgacatcagagctttaaattggtgagcatatcagagacat 1238
Db 3125 GGTGAGTGTATGAGACAGATTAATGAGACCATTAATGTAAGTTTCTACGGGAAATCAT 3184
QY 1239 tgaagtgatgaagaagcctgttcttgcggagacacacatcagtaacatcaatga 1298
Db 3185 TGCTGTGTGATGACGAGAACCTGATTTGTTGCCACGACGATGAAACATTCGCTA 3244
QY 1299 tgaacgagatgattgactgataagagatgagagagcagcaaggaagaacaaatgct 1358
Db 3245 TGCCCGTGAATATGTCACCATGATGAGATTGAGAAAAGCTGTCAAGGAAGCAATGCTTA 3304
QY 1359 tgatttatacttgagattcctaataaatttaacatttgtagggaaagagctca 1418
Db 3305 TGACTTTATCATGAAATGCTCATTAATTTGACACCTGTTGAGAGAGAGGGGCCA 3364
QY 1419 aatgattgagagcaagaacagagatgcgaattgctcgttctttagtgcgaacccca 1478
Db 3365 GTTAGAGTGTGGGACGAAGCAGAGATGCGCATTTGACAGTCCCTGTTGCCAACCCAA 3424
QY 1479 gattctgatttagatgagctacgctgccttgatctagaagaagcaagtcagctgtca 1538
Db 3425 GATCCTCTGCTGTGATGAGGCGACGTCAGCTTGACACAGAAAGCGAAGCATGTTCA 3484
QY 1539 agctgcactgagaag 1555
Db 3485 GGTGGCTGTGATTAAG 3501

```

Search completed: April 22, 2002, 22:55:48  
 Job time: 27856 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 23:10:39 ; Search time 886.72 Seconds

(Without alignments)  
1875.689 Million cell updates/sec

Title: US-09-873-409-16

Perfect score: 1940

Sequence: 1 ttccgcttgcgtatgact.....taaatgcgacttatgta 1940

Scoring table: IDENTITY\_NDC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N.Geneseq.1101.\*

1:	/SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	624.6	32.2	4279	22 AAD03506
2	623	32.1	4279	22 AAD03504
3	623	32.1	4279	22 AAD03505
4	621.4	32.0	4279	22 AAD03488
5	619.8	31.9	4317	22 AAD03489
6	603.8	31.1	3860	21 AAZ49332
7	603.8	31.1	4349	22 AAH57442
8	603.8	31.1	4646	15 AAQ72872
9	603.8	31.1	4669	19 AAV32645
10	603.8	31.1	4669	19 AAV32645
11	603.8	31.1	6505	17 AAT13394

12	603.8	31.1	8630	21 AA224042	Retroviral vector
13	602.2	31.0	3860	21 AA249333	Human G185V mutant
14	602.2	31.0	4669	8 AAN70752	Sequence of human
15	602.2	31.0	4669	14 AAQ52726	Sequence of human
16	602.2	31.0	8630	21 AA224041	Retroviral M4 mdr-
17	600.6	31.0	4378	11 AAQ04522	Multidrug Resistant
18	595.8	30.7	4186	22 AAF86127	Cynomolgous monkey
19	595.8	30.7	4186	22 AAF86128	Cynomolgous monkey
20	594.4	30.6	4264	19 AAV65533	Mutated human P-g1
21	592.8	30.6	4264	19 AAV65534	Mutated human P-g1
22	589.4	30.4	3988	21 AA288973	Human MDR-1 DNA.
23	571.6	29.5	4189	21 AA249334	Murine multidrug r
24	571.6	29.5	4313	14 AAQ38950	Mouse multidrug r
25	568.4	29.3	4233	21 AA290198	Rat mdr1b2 (multis
26	568.4	29.3	4233	22 AAF27498	Rat mdr1b2 multidr
27	568.4	29.3	4788	21 AA249335	Murine multidrug r
28	560.2	28.9	4425	21 AA252048	Rat multidrug resi
29	558.6	28.8	4369	21 AA252047	Rat multidrug resi
30	536.4	27.6	3924	21 AA294742	Human ATP binding
31	536.4	27.6	3924	21 AA288974	Human MDR-3 DNA.
32	439.4	22.6	4776	21 AA294744	Human ATP binding
33	300.6	12.5	4102	21 AAC45942	Arabidopsis thalia
34	297.2	15.3	1810	20 AAV69393	H. contortus PGP-A
35	297.2	15.3	4175	20 AAV69392	H. contortus PGP-A
36	258	13.3	2726	15 AAQ70907	Multidrug-resistant
37	258	13.3	2726	15 AAQ70916	Multidrug-resistant
38	258	13.3	2726	18 AAT43322	Multidrug resistant
39	248.6	12.8	4051	21 AAA09019	A. thaliana ATPAC
40	243.2	12.5	4002	21 AAX85823	cDNA encoding mult
41	238.6	12.3	2779	22 AAS22445	Human cDNA encodin
42	230.8	11.9	4224	19 AAV42347	cDNA encoding a mu
43	219.4	11.3	4047	20 AAX60201	cDNA encoding the m
44	219.4	11.3	4800	20 AAX60202	DNA encoding the m
45	217.8	11.2	4414	21 AAC47124	Arabidopsis thalia

## ALIGNMENTS

RESULT 1	
AAD03506	
ID	AAD03506 standard; cDNA; 4279 BP.
XX	
AC	AAD03506;
XX	
DT	13-JUN-2001 (first entry)
XX	
DE	Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX	
KW	Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW	MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX	
OS	Canis familiaris.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	17..3862
FT	/*tag= a
FT	/product= "Dog P-glycoprotein (PGP) allelic variant
FT	(Genotype D) protein"
FT	replace (91, T)
FT	/*tag= b
FT	replace (607, C)
FT	/*tag= c
FT	replace (1001, T)
FT	/*tag= c
FT	replace (3458, A)
FT	/*tag= c
PN	WO200123540-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26767.

XX 28-SEP-1999; 99US-0156510.  
PR  
XX  
PA (GENT-) GENTEST CORP.  
PI  
XX Stocker PJ, Steinmetz DT, Crespi CL, Reif TC, Patten CJ;  
DR WPL; 2001-235373/24.  
XX P-PSDB; AAE00310.  
PT  
XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
XX PGP inhibitors -  
XX  
XX Claim 9; Page 102-107; 111pp; English.

The invention relates to dog P-glycoprotein (pgp) also referred as multidrug transporter (MDR) and nucleic acids encoding them. The invention also includes fragments and biologically functional variants of dog P-glycoprotein. Pgp and their nucleic acids are useful for determining the bioavailability of drugs and for screening PGP inhibitors. They are useful for the diagnosis and treatment of conditions characterised by PGP activity, by reducing or increasing PGP activity in a cell. PGP nucleic acids are used as oligonucleotide probes. Complements of PGP nucleic acids are useful as antisense oligonucleotides, to induce a PGP 'knockout' phenotype. They are used to prepare a non-human transgenic animal, which are valuable as genetic models for human diseases.

The present sequence is dog P-glycoprotein (PGP) allelic variant (genotype D) cDNA. The PGP enzyme functions as an efflux pump exporting small molecules across the cell membrane. This enzyme is a member of the ABC transporter family.

Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Query Match	32.2%	Score	624.6	DB	22	Length	4279
Best Local Similarity	66.5%	Pred. No.	4.6e-177				
Matches	929	Conservative	0	Mismatches	459	Indels	9
						Gaps	2

QY	162	caggctgacctgattatgtgtggaatgaatgtgtgccttgattttgttgaatcaca	221
	163		
	164		
	165		
	166		
	167		
	168		
	169		
	170		
	171		
	172		
	173		
	174		
	175		
	176		
	177		
	178		
	179		
	180		
	181		
	182		
	183		
	184		
	185		
	186		
	187		
	188		
	189		
	190		
	191		
	192		
	193		
	194		
	195		
	196		
	197		
	198		
	199		
	200		
	201		
	202		
	203		
	204		
	205		
	206		
	207		
	208		
	209		
	210		
	211		
	212		
	213		
	214		
	215		
	216		
	217		
	218		
	219		
	220		
	221		
	222		
	223		
	224		
	225		
	226		
	227		
	228		
	229		
	230		
	231		
	232		
	233		
	234		
	235		
	236		
	237		
	238		
	239		
	240		
	241		
	242		
	243		
	244		
	245		
	246		
	247		
	248		
	249		
	250		
	251		
	252		
	253		
	254		
	255		
	256		
	257		
	258		
	259		
	260		
	261		
	262		
	263		
	264		
	265		
	266		
	267		
	268		
	269		
	270		
	271		
	272		
	273		
	274		
	275		
	276		
	277		
	278		
	279		
	280		
	281		
	282		
	283		
	284		
	285		
	286		
	287		
	288		
	289		
	290		
	291		
	292		
	293		
	294		
	295		
	296		
	297		
	298		
	299		
	300		
	301		

OY	639	agaactcaaaagatatacacaagaatctcaaaagatcaaaaggttttgctcataaaagaagac	698
Db	838	agaacttgaagaagatgacaacaanaaattcagaagaagcacaaggaatttggatacaagaagaac	897
OY	639	tatagtctcaaaagtgtcctcttgtagctgtgtaacttcttiaagaatggaactatgtaact	758
Db	898	tatacagcgcaacttcttatgttgtagcgtcttcttatgtatcatatgatacatatgctct	957
OY	759	tgcttttttgtagtagaacctctcttgattcttaatgagaacctgtgatataccatcggaac	818
OY	819	tgctcttgtagctttctcttgtagttaaatacagaagtaatgtacatcttgtagagcaagctccc	878
Db	1012	agtaaccacgtctctctctctctgtatctaatttggcgttttagtatattggaacgaatcccc	1071
OY	879	tcaacttgaacaaccttcgcaataagcccgaggaagctgcctcttaactatcttcacagttatgta	938
Db	1072	aagatcttgaagaacttttgcaaacgcaagaggagcagcttatgtaatacttccaagatatgta	1131
OY	939	taagaaacccagtatatagataactttccacagctgtgatataaactgtgaatccaatagaag	998
Db	1132	caataaaccacaagatattacacgctattcgcgaagatgtagacaataacagataataactaaagg	1191
OY	999	aactgtgaatttaaaaatgtttcttccaattatccatacaagaacatctacatgaactatct	1058
Db	1192	aaatttggaaattcaaaaatgtaatttcaacttccatgtaaccttctcgaaaagaagttaagaacct	1251
OY	1059	gaaaagctctaatctcgcgaatttaagctctgtagagaagacagctgccttgctgcgtctcaatg	1118
Db	1252	aaagggtctcaacctggaagttcagaagtgtagagacagatgagctgtctgtaggaacagtg	1311
OY	1119	cagtgtggaagaagatgacgtgtagtcacgcttctcgcaagagtataatgataccgagtatgacct	1178
Db	1312	ctcgcggaagaagacagacgacgcgtgcagctgtatgacagaaggtctctatgacccacaagatgagat	1371
OY	1179	tatatagtgtagtagagatatacactcgaagcttaaatgttgtagagcaattatcgagaacatat	1238
Db	1372	ggctctgattctgtagagacagagaacattagagacataatgtaaggaacacctcttcgggaataatc	1431
OY	1239	tgaagtggttagtcaagaagagctgttttctgtcggagacacacatcaaglaatacaatatacaagta	1298
Db	1432	tggtgtgtgtgtagtcaagagagagcgtgtgttgttgcacacagatagtagtaaaacattgccta	1491
OY	1299	tgaagacagatgtagtgcactgtatgaagaagatgtagagaagcagacgaagtggaagcaaatgccta	1358
Db	1492	tgagcgcggaaatgtctaccctatgtagatgtgaagaagcgtcttaagagaagcaaatgtgccta	1551
OY	1359	tgattttatcatatgagttcttcctcttaaatattatcatatcatgttaggtaggaagaagagctca	1418
Db	1552	tgattttatcatatgaaactactctaataatttttaacaactctgtgttgtagagagagaagggccca	1611
OY	1419	aatgaagtgaaggcgacagaacaagagatactgcgaattgctgtgcttagtctgaacacccca	1478
Db	1612	gtctgtagtggtagacagaagaagaagaaatgcgaattgctgtcgagccctgtgtgcgaacccca	1671
OY	1479	gattctgattttagatgtaggctacgtctgcgcctgtgattcagaagaagcaagtcatgctgttca	1538
Db	1672	gattcttctctgtagatgtagcaagctcagctctgtgaacactgaaagtgaagacagtggttca	1731
OY	1539	agctgcacgtgagagaag	1555
Db	1732	ggtagccctgtgataagg	1748

RESULT	2	
AAD03504		
ID	AAD03504	Standard; CDNA; 4279 BP
XX		
AC	AAD03504;	
XX		
DT	13-JUN-2001	(first entry)
XX		

DE Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.  
 XX  
 XX Dog: P-glycoprotein allelic variant; multidrug transporter; MDR1;  
 KM drug bioavailability; transgenic animal; genetic model; ss.  
 XX  
 OS Canis familiaris.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 17..3862  
 FT /\*tag= a  
 FT /product= "Dog P-glycoprotein (PGP) allelic variant  
 FT (Genotype A) protein"  
 FT replace (607, C)  
 FT allele /\*tag= b  
 XX  
 XX WO200123540-A2.  
 XX  
 XX 05-APR-2001.  
 XX  
 XX 28-SEP-2000; 2000MO-US26767.  
 XX  
 XX 28-SEP-1999; 99US-0156510.  
 XX  
 XX (GENT-) GENTEST CORP.  
 XX  
 XX Stocker PJ, Steimel-crespi DT, Crespi CL, Relif TC, Patten CJ;  
 PI WPI; 2001-235373/24.  
 DR P-PSDB; AAE00308.  
 XX  
 XX New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
 PT for determining the bioavailability of drugs and for screening for dog  
 PT PGP inhibitors -  
 XX  
 XX Claim 9; Page 85-90; 11pp; English.  
 XX  
 CC The invention relates to dog P-glycoprotein (PGP) also referred  
 CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
 CC The invention also includes fragments and biologically functional  
 CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
 CC useful for determining the bioavailability of drugs and for  
 CC screening PGP inhibitors. They are useful for the diagnosis and  
 CC treatment of conditions characterised by PGP activity, by  
 CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
 CC are used as oligonucleotide probes. Complements of PGP nucleic  
 CC acids are useful as antisense oligonucleotides, to induce a PGP  
 CC 'knockout' phenotype. They are used to prepare a non-human  
 CC transgenic animal, which are valuable as genetic models for  
 CC human diseases.  
 CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
 CC (Genotype A) cDNA. The PGP enzyme functions as an efflux pump  
 CC exporting small molecules across the cell membrane. This enzyme  
 CC is a member of the ABC transporter family.  
 CC  
 XX Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Query Match 32.1%; Score 623; DB 22; Length 4279;  
 Best Local Similarity 66.4%; Pred. No. 1.4e-176;  
 Matches 928; Conservative 0; Mismatches 460; Indels 9; Gaps 2;

QY 162 caagttgacccgtatattatgttgaatagtgctgtgctgtgattttgtgtacataca 221  
 DB 358 cactgtatgctattatctatgagtgatcggtgtgctgtgctgtgctgtgtacataca 417  
 QY 222 gattcttggtgattatacttcacgacacgacacgaaggatctggaagaattttt 281  
 DB 418 ggttcaattctgtgctgtgacgacgaagaaacgatactaaattagaacaattttt 477  
 QY 282 tcatcagtttctgacgacacacgacacgctgtgtgttgaatgctgtgacatcgttgaactaa 341  
 DB 478 tcatgcatcatatgacgacgagatgtgctgtgttgaatgcatgacgttggggagactaa 537

QY 342 cactgcacatgacagacattg----acaaaatcagtatgtgtattgagatcaaatgtcct 398  
 DB 538 caccgcgtccacagacagatgctccaaaatcgaatgaaggaatttggacaaaatttgaat 597  
 QY 399 gtctttcaaaaacatgcttactcttcttgaattgctgtgacgttgtgttggtaaggcttg 458  
 DB 598 gtctttcaatcaatgaacaaattttcaaccggttttattagtggtgtttcaacgtgttg 657  
 QY 459 gaaactcaacctgagtgactctatccagctcctccttataatgcttcaaggcgcacatg 518  
 DB 658 gaagctcaaacctgtgtatttcttgcacacgacctgttcttggacttccgcacatcg 717  
 QY 519 ttctagatggtatcctatctgacccagtaaggatgaattgaatgcttccaaagcttgggc 578  
 DB 718 ggcacaaagatactatcttcaattactgtataagaaccttggccttgaacaaagcttgagc 777  
 QY 579 tgtgacagaagaatgcttctgtatcaatccgaaacagtcataagctttagggcccaagaga 638  
 DB 778 agtagctgaagaatgcttctgtacagcaatcagaactgtgatgtgcttggagacaaagaa 837  
 QY 639 agaaactcaaaaggtatatacacaagaatctcaagaatgcaaaagatttggcataaaagac 698  
 DB 838 agaaacttgaaggtatatacacaagaatctcaagaatgcaaaagatttggcataaaagac 897  
 QY 699 tatagttcaaaaagtgtctctgtgtgtgtgtgtatcttcttgaatggaacctatgagact 758  
 DB 898 tatcacgaccaaattctcatgt 957  
 QY 759 tgcctttgtatggaacccctctgtatcttcttaatggaagacatgatatccatccggagac 818  
 DB 958 ggccttctgtatggaacccctctgtatcttcttaatggaagacatgatatccatccggagac 1011  
 QY 819 tgccttctgtatggaacccctctgtatcttcttaatggaagacatgatatccatccggagac 878  
 DB 1012 agtactcactgactcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1071  
 QY 879 tcaacttgaacacttctgcaatagccggaagagctgcttcttcttcttcttcttcttcttct 938  
 DB 1072 aagacttgaacacttctgcaatagccggaagagctgcttcttcttcttcttcttcttcttct 1131  
 QY 939 taagaacccacagatagatatacttctcacaagcttgaatataaaacttgaatccatagaag 998  
 DB 1132 caataaacccaagatagatatacttctcacaagcttgaatataaaacttgaatccatagaag 1191  
 QY 999 aactgtggaatttaaaaatgttcttcttcaatttccatcaagaagccatctcaagatct 1058  
 DB 1192 aaatttgaatttcaaaaatgttcttcttcaatttccatcaagaagccatctcaagatct 1251  
 QY 1059 gaaaggtctgaatctcaagaatgaattgaattgaagagacagctgtgtgtgtgtgtgtgt 1118  
 DB 1252 aaaggtctgaatctcaagaatgaattgaattgaagagacagctgtgtgtgtgtgtgtgtgt 1311  
 QY 1119 caatggaagaagatgaattgaattgaattgaattgaattgaattgaattgaattgaattga 1178  
 DB 1312 ctgctggaagaagatgaattgaattgaattgaattgaattgaattgaattgaattgaattga 1371  
 QY 1179 tatcatgtgtgattgagatgagatgacatcagaagcttcaaatgtgcgcatatcagaacat 1238  
 DB 1372 gctcgtgattgagatgagatgacatcagaagcttcaaatgtgcgcatatcagaacat 1431  
 QY 1239 tgaagtgtgtgattgagatgagatgacatcagaagcttcaaatgtgcgcatatcagaacat 1298  
 DB 1432 tgaagtgtgtgattgagatgagatgacatcagaagcttcaaatgtgcgcatatcagaacat 1491  
 QY 1299 tgaagtgtgtgattgagatgagatgacatcagaagcttcaaatgtgcgcatatcagaacat 1358  
 DB 1492 tgaagtgtgtgattgagatgagatgacatcagaagcttcaaatgtgcgcatatcagaacat 1551  
 QY 1359 tgaatttataatgagatttccataataatcagaatgagatgagatgagatgagatgagatg 1418  
 DB 1552 tgaatttataatgagatttccataataatcagaatgagatgagatgagatgagatgagatg 1611  
 QY 1419 aatgagtgagagagacagaagaagatgagatgagatgagatgagatgagatgagatgagatg 1478

||||| || ||||||| ||||| ||||||| || ||||| |||||||  
Db 1612 gctgagtggtgacagaacagagatgcattgctcgccctggtcgaacccaa 1671  
Oy 1479 gatctgatttaagatgagtgctgctgctgattcagaagaagtacgtttca 1538  
||||| || ||||||| ||||| ||||| ||||| ||||| |||||  
Db 1672 gatcttctgctgagtgaagcaagtcagctctgacactgaagtgaaagtggtca 1731  
Oy 1539 agctcaactgagaag 1555  
| || ||||| |||||  
Db 1732 ggtgacctgataag 1748

RESULT 3  
AAD03505  
ID AAD03505 standard; cDNA: 4279 BP.  
AC AAD03505;  
XX  
DT 13-JUN-2001 (first entry)  
XX  
DE Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.  
XX  
KW Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;  
KM MDR1; drug bioavailability; transgenic animal; genetic model; ss.  
XX  
OS Canis familiaris.  
XX  
FH Key Location/Qualifiers  
FT CDS 17..3862  
FT /\*tag= a  
FT /product= "Dog P-glycoprotein (PGP) allelic variant  
FT (Genotype B) protein"  
FT replace (91, T)  
FT /\*tag= b  
FT replace (607, C)  
FT /\*tag= c

MO200123540-A2.  
XX  
PN 05-APR-2001.  
XX  
PD 28-SEP-2000; 2000MO-US26767.  
XX  
PF 28-SEP-1999; 99US-0156510.  
XX  
PR 28-SEP-1999;  
XX  
PA (GENT-) GENTEST CORP.  
XX  
PI Stocker PJ, Stelmal-crespi DT, Crespi CL, Reif TC, Patten CJ;  
XX  
DR WPI: 2001-235373/24.  
DR P-PSDB; AAE00309.  
XX  
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful  
PT for determining the bioavailability of drugs and for screening for dog  
PT PGP inhibitors -  
XX  
PS Claim 9; Page 93-99; 111pp. English.

CC The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity. By  
CC reducing or increasing PGP activity in a cell, PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.  
CC The present sequence is dog P-glycoprotein (PGP) allelic variant  
CC (Genotype B) cDNA. The PGP enzyme functions as an efflux pump

CC exporting small molecules across the cell membrane. This enzyme  
CC is a member of the ABC transporter family.  
XX  
SQ Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Query Match 32.1%; Score 623; DB 22; Length 4279;  
Best Local Similarity 66.4%; Pred. No. 1.4e-176;  
Matches 928; Conservative 0; Mismatches 460; Indels 9; Gaps 2;

Oy 162 caggtgacctgtattatgttgatagtggtgctgcttgattttgtttacataca 221  
||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 358 cagatgacctattattacagtgagatgctgctgctgctgctgctgctgctgct 417  
Oy 222 gatcttctggtgattataactgcagcagcagcagcagcagcagcagcagcagc 281  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 418 ggttcattctggtgctgctgctgctgctgctgctgctgctgctgctgctgct 477  
Oy 282 tcatcaatttggcacagacatcgctggttgatagctgtgacatcggtgacattaa 341  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 478 tcatgctatcatgacagagagatgctgctgctgctgctgctgctgctgctgct 537  
Oy 342 cactgcgatgacagacatg---acaaaatcagatgattgttgagataagattgct 398  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 538 caccggctcacagacgagtctccaaaataatgaagaattggtccaaaattggaat 597  
Oy 399 gttgttcaaaaacatgctactcttctgattgctgctgctgctgctgctgctgct 458  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 598 gtcttccaatcaatagaacaattttccagtggtttatagtggtttacacggtgtg 657  
Oy 459 gaaactcaccctagtgactctatccagctctctcttataatgcttcaagcgacatg 518  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 658 gaagctaaccttgatttgctgcatcagccctgtcttgacttcaagcgcacatgt 717  
Oy 519 ttctagatgtcatcctaatgagcaagtaagatgacctatccatccaaagctgggg 578  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 718 ggcagaagatacttcatcttaactgaagaagaactcttgctgaagaagctggagc 777  
Oy 579 tgtgcagaagaagctctgtcatccatccagacagatcagaccttaaggccagaaga 638  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 778 agtagctgaagaagctcttagcagcatcagacatgctgcttcttgtagagcaaaaga 837  
Oy 639 agaacttcaaaagttatcacagaaatcacaagatgcaagatttgcataaaaaggac 698  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 838 agaactgaaggtacacaaaatttagaagaagctaaagatggtgataaagaagc 897  
Oy 699 tatagttcaaaagtgctctgtgctgctgctgcttcttataatgagtaagtaagact 758  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 898 tatcagggccaaacttctatgtgtgctgcttcttattgatactatgatacatatgctct 957  
Oy 759 tgccttggatggaacctctcttattcttataatgagaacctgatatcatcccgagac 818  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 958 ggccttctggtatggacctctctgtctctccg-----tgataattcatggaca 1011  
Oy 819 tgttctgctgttctctttagtgaatccatgacagattatgcatgagcagctccc 878  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1012 agtactcaactgctctcttcttctgatttaattgaggtctttagttagcaagcatccc 1071  
Oy 879 tcaacttgaaccttcgcaatagccggagagctgctcttcatatttccaggttatga 938  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1072 aagcatgaagatttgaagaagcaagagagcagcttatgaaattccaagataatgca 1131  
Oy 939 taagaaccagatagataaacttccacagctgatatataaactgataatcagagaag 998  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1132 caataaaccagatgacagctatctgaagagtgagataaaccagataaataaagag 1191  
Oy 999 aactgtgaatttaaaatggttcttcaatataccatcaagaacatctatcaagattct 1058  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1192 aaatttgaattcaaaaatggttcttcaatgattacccctctcgaagaagttaaagctt 1251  
Oy 1059 gaaaggtctgaatctccagaatgaatctggaagaagcagctgctgctgctgctaaagg 1118  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1252 aaaggtctcaacctgaaggttcaagagtggtgacagacatggtgctgtgttggaacaagtg 1311

OY	1119	caatgggaaagatcagtaatccaagctcttcgaagaggtatataatgtaaccgatgatgctt	1178
Db	1312	ctggggaaagaagccagcccttgagcttgatgcagaggtctctatgaccaccaagatgcat	1371
OY	1179	tatcatgttgagatgagatcacaaacgaagcttaaaatgtagcgacatttcgaagccatat	1238
Db	1372	ggtcttgatcttgatgagaagaagcaactctgagaccataaattgtaagagctctcttggaaattac	1431
OY	1239	tggagtggaatgaatgaagaagcctgttttgcctgcgagaccacatcaatlaacatlaagta	1298
Db	1432	tgtgtgtgtagatcaagaagagcctgtgtgtgttcgcccaacgatactgtgaaacattgccta	1491
OY	1299	tggagcgaatgagatgtgctctgataaagaatgtgaagagagaagaagaagcaaatgtcta	1358
Db	1492	tggcgcgcgaaaaatgaccaccaatggaatggaagaagcgtgttaaggaagccaatgctta	1551
OY	1359	tgaattatcatgtgaagttcttccataaactatatacatgtgtgaagggaagaaagagctca	1418
Db	1552	tgaattatcatcaactactactaataatttgaacactctggtgtggaagagaggggccca	1611
OY	1419	aataagtgaggggcgaagaagaagaagatctgcgaattgctgtgccttaagtccaaccccaa	1478
Db	1612	gcctgaagtgagtgagaagaagaacaagaagaatcgccatctgcgtggccctggtctgcacacccaa	1671
OY	1479	gattctatattatgatatgagctacgcctcccgatgttaagaagaagaatcagatgcttca	1538
Db	1672	gattctctctctgtgatgagggcaacgctcagctctcgaacactgtgaagaacgtgaacagtggttca	1731
OY	1539	agctcgcaactgaggaag	1555
Db	1732	ggttgcccttgataaag	1748

XX	RESULT	4
AD03488	AD03488	standard; cDNA; 4279 BP.
AC	AD03488;	
XX	13-JUN-2001	(first entry)
DE	Dog P-glycoprotein (PGP)	CDNA #1.
XX	Dog: P-glycoprotein; PGP; multidrug transporter; MDR1;	
KW	drug bioavailability; transgenic animal; genetic model; ss.	
XX	Canis familiaris.	
OS		
XX	Key	Location/Qualifiers
FH	CDS	17..3862
FT		/*tag= a
FT		/product= "Dog P-glycoprotein (PGP) #1"
XX		
PN	MO200123540-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000MO-US26767.	
XX		
PR	28-SEP-1999; 99US-0156510.	
XX		
PA	(GENT-) GENTEST CORP.	
XX		
PI	Stocker PJ, Steinmel-crespi DT, Crespi CL, Reif TC, Patten CJ;	
XX		
DR	WPI; 2001-235373/24.	
DR	P-PSDB; AAEO0303.	
XX		
PT	New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful	
PT	for determining the bioavailability of drugs and for screening for dog	
XX	PGP inhibitors -	

PS Claim 5, Page 58-63, 11pp; English.

XX

CC The invention relates to dog P-glycoprotein (PGP) also referred  
CC as multidrug transporter (MDR1) and nucleic acids encoding them.  
CC The invention also includes fragments and biologically functional  
CC variants of dog P-glycoprotein. PGP and their nucleic acids are  
CC useful for determining the bioavailability of drugs and for  
CC screening PGP inhibitors. They are useful for the diagnosis and  
CC treatment of conditions characterised by PGP activity, by  
CC reducing or increasing PGP activity in a cell. PGP nucleic acids  
CC are used as oligonucleotide probes. Complements of PGP nucleic  
CC acids are useful as antisense oligonucleotides, to induce a PGP  
CC 'knockout' phenotype. They are used to prepare a non-human  
CC transgenic animal, which are valuable as genetic models for  
CC human diseases.

CC The present sequence is dog P-glycoprotein (PGP) cDNA. This  
CC sequence is also referred as Genotype C cDNA. The  
CC PGP enzyme functions as an efflux pump exporting small molecules  
CC across the cell membrane. This enzyme is a member of the ABC  
CC transporter family.

XX

SQ Sequence 4279 BP: 1294 A; 834 C; 1008 G; 1143 T; 0 other;

[illegible]

```
Db 1012 agtactcactgtctctcttcttcttaatttgggctttagtattgagcagcatcccc 1071
Oy 879 tcaacttgaaccttcgaactagccccggagctgcttcttcttatttccagtatttga 938
Db 1072 aagcattgaagatttgcacaacgcaagagagcattatgaacttctcaagatatga 1131
Oy 939 taagaacacagatagataaacttccacagctgataataaactgaattcagagaag 998
Db 1132 caataaaccacagcattgcagctattcgaaagtgagactaaacagataataatgaag 1191
Oy 999 aactgtggaatttaaaaatgttcttcttaactatccatcaagaacatctataaagattc 1058
Db 1192 aaatttgggaattcaaaaatgttcaacttcaagttaccccttcgaaaaagaagtaagatct 1251
Oy 1059 gaaagctgaactcagaaattagttcggaagagacagtcgcttggtgcgtctaaag 1118
Db 1252 aaagtgctcaacccggaaggttcagagtcgagacagtgctggttggaacagtg 1311
Oy 1119 cagtggaagagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1178
Db 1312 ctgcggaagagagagacagcagctgagctgagagagctcttgacccacagatgagct 1371
Oy 1179 tatcagtgatgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1238
Db 1372 gctcgtatgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1431
Oy 1239 tggagtgttgtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1298
Db 1432 tgggtgtgtgtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1491
Oy 1299 tggaggaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1358
Db 1492 tggcgcggaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1511
Oy 1359 tgatttcatgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1418
Db 1552 tgatttcatgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1611
Oy 1419 aatgagtgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1478
Db 1612 gctgagtgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1671
Oy 1479 gatctgatttgaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1538
Db 1672 gatcttctgcgagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 1731
Oy 1539 agctgcactgagaag 1555
Db 1732 ggtggtccttgataag 1748

RESULT 5
AAD03489
ID AAD03489 standard; cDNA: 4317 BP.
AC
AAD03489:
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #2.
XX
KW Dog: P-glycoprotein; PGP: multidrug transporter; MDR1:
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 70..3912
FT /tag=a
FT /product= "Dog P-glycoprotein (PGP) #2"
XX
PN W0200123540-A2.
```

```
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI: 2001-235373/24.
XX
P-PSDB: AAE00304.
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX
PS Claim 1; Page 66-72; 11pp; English.
XX
CC The invention relates to dog P-glycoprotein (PGP) also referred
CC as multidrug transporter (MDR1) and nucleic acids encoding them.
CC The invention also includes fragments and biologically functional
CC variants of dog P-glycoprotein. PGP and their nucleic acids are
CC useful for determining the bioavailability of drugs and for
CC screening PGP inhibitors. They are useful for the diagnosis and
CC treatment of conditions characterised by PGP activity, by
CC reducing or increasing PGP activity in a cell. PGP nucleic acids
CC are used as oligonucleotide probes. Complements of PGP nucleic
CC acids are useful as antisense oligonucleotides, to induce a PGP
CC 'knockout' phenotype. They are used to prepare a non-human
CC transgenic animal, which are valuable as genetic models for
CC human diseases.
CC The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC PGP enzyme functions as an efflux pump exporting small molecules
CC across the cell membrane. This enzyme is a member of the ABC
CC transporter family.
XX
SQ Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Query Match 31.9%; Score 619.8; DB 22; Length 4317;
Best Local Similarity 66.3%; Pred. No. 1.3e-175;
Matches 926; Conservative 0; Mismatches 462; Indels 9; Gaps 2;

Oy 162 caggttgaccctgtattatgttggaatagtggtgtctgtctgtatttggttacata 221
Db 408 caggtatgctattattacagtgagatcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 467
Oy 222 gatttcctgtgattataactgcagcagcagcagcagcagcagcagcagcagcagcagc 281
Db 468 gtttcattctgtgtcctgcgcagcagcagcagcagcagcagcagcagcagcagcagc 527
Oy 282 tcaatcagtttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 341
Db 528 tcatgtatcatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 587
Oy 342 cactcgcagatgacagatc---acaaatcagtgatgtatgtatgtatgaatattgctct 398
Db 588 caccgcgtcacagcagatgtctccaatcagtaagtaagtaagtaagtaagtaagtaag 647
Oy 399 gtgtttcaaaacatgtacttttgcagtcgctgcagtcagtcagtcagtcagtcagtcag 458
Db 648 gtctttcaataaagaacatttttccacggtttattatagtggtgttataacacctgtgt 707
Oy 459 gaaactcacccatgactctacacagctctccttataatgatgctcagcgcagcagcagc 518
Db 708 gaagtaacccttgatgttttgcacatcagccctgttcttggacttcaagccgcagctgt 767
Oy 519 ttctagatgtcatcatcattgacgaatgaagaatgaatgaatgaatgaatgaatgaatga 578
Db 768 ggcagaatatactatcttattcattcagtaagaactcttggtgcctatgacaaagcgtgagc 827
```



QY	579	tgtygcgaagaagaatcttctgtcatcaatccgaacaagfclatagccttttaagggcccaaggagaa	638
Db	828	agtagcttgaagaagctcttgagcaagaatcagaaacttgatctgtcttggagagcaaaagaa	887
QY	639	agaacttcaagtgatcacacgaactcccaagatgcaagatgcaaaagatcttggcataaaagac	698
Db	888	agaacttgaagaagatcacacaanaaaattagaagaagcttaaaagaatttggatataagaagc	947
QY	659	catagctcaaaagtgctccttggtgtgctgtgtacattctctatgatatgatggaacctatgact	758
Db	948	tatcagcgccacaactctcatatgtgtgcgcttctctattgatctatgatcatcatatgctct	1007
QY	759	tgcttttggaatgaagaccccttgatcttaatgtagaacccgtgatatcatcatctggagc	818
Db	1008	ggcttcctcggaatgagggacctctctgtgcccctccag-----tgaatatcatcttgaca	1067
QY	819	tgcttcttgctttctctttagtvgtaatccataagatattatgcatgtggagcagaatccc	878
Db	1062	ggtactcaactgtctctcttcttgatataattvggctttagtattggaaagcagccccc	1122
QY	879	tcaatttgaaaaccttcgcaatagccgagagctgctcttcatatatttccaggtatgtga	938
Db	1122	aagcatttgaagcatttgcnaacgcaagaaggagcagcttgaatacttcaagaataatga	1183
QY	939	taagaaccccgatagataacttccatccagctctgatatataaacctggaatcccatagaag	998
Db	1182	caataaaccgaagcatgtgacagctattccgaagagtgtgcaataaaccggaataataattaag	1243
QY	999	aacttgtgaatttaaaatglttctcttcaatatlccatacagaaccatctatcacagaattct	1056
Db	1242	aaatttgaattccaanaaatgttccactcagttacccctctcggaaagaagttaagctct	1307
QY	1059	gaagaagcttgaattctcgaatatgaatgtctggaagagacagctgcctgtgtcggtctcaatg	1111
Db	1302	aaaggtgttccaaccttgaaagtgtccagaatgtggcagacagatggtgcgtgttgggaacagtgtg	1367
QY	1119	caagtggaaagatgcggttagtccagcttccgcaagagttaatgatccggatgatagctgt	1176
Db	1362	ctgcgaggaagcagcagccgctgtgcagctgtatgcagaggtcttatgaccaccaagatgcat	1422
QY	1179	tatcatgttggaatgagaatgacatcagagctttaaatgtgtgcgcatlatacagaccatat	1233
Db	1422	ggtctgtatctgatgacagagcatltagaccataaagaatgaagcatctctcggaaattacc	1483
QY	1239	tgtagtgtttagttaaagaagcgtgttctgttcggagccacatcagtaacaatacaagta	1299
Db	1482	tgggtgtgtgtgcacagagccgtgtgtgttgcacacaagatgcctgaaaaactctgccta	1543
QY	1299	tggacgagaatgatgtgcctgaatgagaagatgagagcagcagaagggaaagcaaatgcgtga	1358
Db	1542	tggccgagcaaaatgtcacocatgagtgatgtggaagaaagcgttaaggaagccaatgccta	1607
QY	1359	tgaatttatcatgtgattctcctaataattatbaatcatatgtgtagtggggaaaaagagctca	1411
Db	1602	tgaatttctatctgaactacataataatttgcacactctgtgtggaagagaagggcccg	1667
QY	1419	aatggttggaaggaagaagaacagaaatcgcgaatttgcgtgcgctttagtttcgnaacccca	1477
Db	1662	gtctggtgtgtgagacaagaacagaaatcgcgaatttgcgtgcggttcttgcgaaccccaa	1722
QY	1479	gattcttgattttagatagagctaaagctgtgccttgatcccttgatccagaagaatctcagctgtta	1536
Db	1722	gattcttctgtgtgataagagcgaagctcagctcctcggaacactgaaagtgaaagcagtggttca	1783
QY	1539	agctgcgaactggagaag 1555	
Db	1782	ggtgtgcctgtgataag 1798	

AC	AAZ49332;
XX	
DT	14-MAR-2000 (first entry)
XX	
DE	Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX	
KM	Multidrug resistance; MDR-1; P-glycoprotein;
KW	transmembrane efflux pump; haematopoietic stem cell; transduction;
KW	bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW	gene therapy; gene replacement; genetic defect; thalassemia;
KW	Gaucher's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
XX	cytokine; wild-type; ds.
OS	Homo sapiens.
FH	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..3843
FT	/tag= a
FT	/product= "Human wild-type MDR-1 protein"
FT	replace (553..555, GTT)
FT	mutation
FT	/tag= b
FT	/note= "cDNA sequence of G185V human mutant MDR-1 given in AAZ49333"
XX	
PN	W09961589-A2.
XX	
PD	02-DEC-1999.
XX	
PF	27-MAY-1999; 99MO-US11825.
XX	
PR	28-MAY-1998; 98US-0086988.
XX	
PA	(SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX	
PI	Sorrentino B, Bunting K;
XX	
DR	MP1: 2000-072615/06.
DR	P-P5DB; AAY58186.
XX	
PT	Ex vivo expansion of hematopoietic stem cells transduced with a
PT	sequence encoding human multidrug resistance-1, used for bone marrow
XX	transplantation -
PS	Claim 10; Page 68-70; 113pp: English.
XX	
CC	This sequence represents cDNA encoding human wild-type
CC	multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC	efflux pump, responsible for the export of drugs from cells,
CC	particularly cancer cells. Wild-type MDR-1 shows increased
CC	resistance to etoposide and decreased resistance to vinca
CC	alkaloids compared with a mutant form (AAY58187) where the Gly at
CC	position 185 is replaced by Val. The invention relates to transducing
CC	haematopoietic stem cells with nucleic acid encoding an MDR protein
CC	and culturing the modified cells. The modified haematopoietic stem
CC	cells are useful in bone marrow transplantation (to reconstitute
CC	haematopoietic systems in patients who have undergone chemotherapy or
CC	radiation therapy) and in ex vivo gene therapy of genetic defects in
CC	cells derived from haematopoietic stem cells, e.g., thalassemia,
CC	Gaucher's disease, sickle cell anaemia or leukaemia. The modified
CC	cells can also be used to identify factors involved in regulating
CC	proliferation and differentiation in haematopoietic stem cells.
CC	Haematopoietic stem cells that express MDR-1 will be protected against
CC	chemotherapeutic agents, so can be engrafted while the patient is
CC	undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC	provides sufficient cells to permit standard biochemical analysis.
CC	Overexpression of MDR-1 allows cytokine-driven expansion of
CC	haematopoietic stem cells by at least 10-fold compared with a maximum
CC	of 4-fold in known procedures.
XX	
SO	Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
XX	

RESULT	6
AAZ49332	
ID	AAZ49332 standard; CDNA; 3860 BP
XX	

**Query Match**

31.18; Score 603.8; DB 21; Length 3860;

Best Local Similarity 65.6%; Pred. No. 7.9e-171;  
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;

```

QY 162 caggttaacctgtatataatgttgaataagtgctgctgtcatttggttacataca 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 336 cagataatgctattatcacagtggaatggtcggggtgctggtctgctcatcca 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 222 gattcctgtggttaactcagcagcagacacacaaaggaggttccaaacgatttt 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 ggtttcaatttggcctgctgagctggaagacaaatacaaaatgaaaaacgatttc 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 tcattcagttttgacagagacacgcgtggtttgataagctgttgacacggtgataa 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 456 tcattcataatgacagagagagataggtgtgttgatgtgacgagtgtggggagcttaa 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 342 caccgcagtcagacagatgac-----aaatcagtgtgtttgagataagattgtcct 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 516 caccgcagttcacaatgatagtctcctaagattaaagaaagtatgtgtgacaaaattggaat 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 399 gtgtttcaaaaacatgtcactttttcattgscctgagctggtttggtggaaggcgtg 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 gtcttttcagtcacatgtggaacatttttcaactggtttatagtagatttacaagtgttg 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 459 gaaactcaacctagtgactatccacgctcctctataatggtctcagcggagacatg 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 636 gaaactcaaccttgattgtgcatcagctcgtctctgactgctcagctgtgtctg 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 519 ttctagatgtgctatcattatgacacagtaagaattaaagtctcatttccaaagctgggc 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 ggcacaaatcattctcattcttacttgaagaacctttagcgtatgtaaaagctggagc 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 579 ttgtgcagaaagaagctctgtcatcatccagacagtcatagcctttagggccagagaa 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 756 agtagctggaagaagctctgtgcgaatgaagaactgtgattcatttgaagacaaaagaa 815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 639 agaactcaaaagtcatacacagaactctcaaaagtcaaaagatttggcataaaaaggac 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 816 agaacttgaagagtcacaacaaaatttagaagaagctaaagaattggtgataaagaagc 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 699 tatagcttcaaaagtgtctctgtgtgtgtgttacttctttagaatgtaactagagact 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 876 tattacagccaatatcttcaatagtgctgtcttccctgtactgtatcttgaattatcttgcct 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 759 tgcctttgtgatagaacctcctgtatcttcaatgaagaacctgataatcacatcgggac 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 936 ggccttcctgatagtgagcacacctgtgtcctcagggga-----attcttatggaca 989
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 819 tgtctgtgctgtttctttagtgaatccatagcagttatgtcatttggagcagacgtccc 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 990 agtactcactgtatcttcttctgtatataatggtgggctttagtgtgagcagcactccc 1049
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 879 tcactttgaaccttcgcagacagagcgtgcttcatatttccaggttatgta 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1050 aagcatggaagcatttgcgaatgcaagagagagcagctatgtaaatcttcaagatattga 1109
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 939 taagaaacccagatagataactttccacagctggaatataaacctgtaaccatagaagg 998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1110 taataagccaagatttgacagctatctgaagagtggtgacaaacagataatattaaagg 1169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 999 aacgtgtgaatttaaaaatgtttcttcaattatccatcaagacacattatcaagattct 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1170 aaatttgaattcagaatgtttcacttcaattaccacatctgaaaagaagttaagatctt 1229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1059 gaaaggtctgaatctcagaataatgaatctgagagagacagctgctggtgtctcaacgg 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1230 gaaaggtccctgaacctggaagtgacagagtggtgagcaggtgtgcccgttggaaaaacgg 1289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1119 cagtggaagagtaacgtgagtcagctcttcagaggtatataatgacgtgagtgatgtt 1178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1290 ctgtgggagagggacaaacagtcagctgtatgacaggtctcatgaccccaagagggtat 1349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1179 tataatgtgtgagaaatgacatcagacagctttaaattgtcggaattatcgagacacat 1238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1350 ggtcagtgatgacagagatattagaccataaatgtlaagtttctacgggaatcat 1409
QY 1239 tggagtggttagtcaagaagcctgtttgttgggacacccctcagtaacaatatcaagta 1298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1410 tgggtgtgtgtaagaagacatgtattgttgcacacagataagctgaaacattcgcta 1469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1299 tggacagatgatgtacatgataagagatgagagagacaaaggaaacaaatgtccta 1358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1470 tggcgtgtgaaaatgtcacacagatgagatgagagaagctgtcaaggagacgaatgcta 1529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1359 tgatttatcatgtaggtttccctaataatlaatacatgtgtaggggaaaaaggagctca 1418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1530 tgacttatcatgtaaacgtccctcataaatttgacacacctgtgtgagagagaggggccca 1589
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1419 aatgagtgaagggaagaagagatgcgaattgtctgtccttgaatgttgcgaaccccaa 1478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1580 gtctgagtggtggcagagacagagatcgcattgcacgtgccccgtgtcgaaccccaa 1649
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1479 gattctgattttagatgaggtacgtctgccttgatbcagaaagcaagtcagctgttca 1538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1650 gatcctcctgtgtgtagagcgccagctcagccttgacacagaaagcgagcagtggttca 1709
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1539 agctgcactgagagaagg 1555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1710 ggtgctctgataaagg 1726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AAH57442
ID AAH57442 standard; cDNA: 4349 BP.
XX
AC AAH57442;
XX
DE 10-SEP-2001 (first entry)
XX
DE Human intestine cell specific cDNA sequence SEQ ID NO:282.
XX
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS Homo sapiens.
XX
PN MO2001.32927-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000OWO-US30396.
XX
PR 04-NOV-1999; 99US-0163508.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sornasse T, Seilhamer JJ, Watson GA;
XX
DR WPI: 2001-291057/30.
XX
PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology -
XX
PS Claim 1; Page 207-208; 327Pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytoskeletal, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by them are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from

```



XX WPI: 1994-332816/41.  
 DR P-PSDB: AAQ72872.  
 XX Liposomes for delivering protein, nucleic acid etc. to hair  
 PT follicles - e.g. to restore hair colour, prevent hair loss during  
 PT chemotherapy, stimulate hair growth etc.  
 XX

XX Claim 8: Page 70-77; 100pp; English.

CC The nucleotide sequence of the human multidrug resistance (MDR-1) gene  
 CC which encodes a P-glycoprotein. This is an example of a compound which  
 CC can be delivered to hair follicles via a novel liposome composition.  
 CC The liposomes are comprised of a lipophilic or lipophobic compound which  
 CC will selectively target the hair follicle (via the stratum corneum)  
 CC without damaging or unwanted effects on cells outside the follicle.  
 CC Compounds e.g. P-glycoprotein, can be delivered to the hair follicle for  
 CC treatment of chemotherapy-induced alopecia. Other compounds targeted at  
 CC hair follicles can include: tyrosinase (or the DNA encoding it -  
 CC AAQ72871), melanin or hair dyes (to restore hair colour or condition);  
 CC human transforming growth factor-alpha (AAQ72873) (for reversal of wavy  
 CC hair); hair growth stimulants (e.g. cyclosporin A or aromatase) or  
 CC antitense sequences. The method allows compounds (e.g. macromolecules  
 CC or polymers), which would not normally reach the hair follicles, to be  
 CC delivered to these target areas.  
 XX

SO Sequence 4646 BP; 1371 A; 893 C; 1128 G; 1254 T; 0 other;

Query Match 31.1%; Score 603.8; DB 15; Length 4646;  
 Best Local Similarity 65.6%; Pred. No. 8.7e-171;  
 Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;

QY 162 caagttgacgctatattgtgaataggtgtgctgcttatttggtaacaca 221  
 DB 760 caggtatcctattatcaatcagatggaattgctgctgctgcttcaatca 819  
 QY 222 gattcccttgatataaactcagacacgacacgaaggaaggttcaaacagtttt 281  
 DB 820 gtttcaatttggctgctgctgctgctgctgctgctgctgctgctgctgct 879  
 QY 282 tcaatcagtttgcacagacacatcgctgttgaagctgtgacatcggtgaactaa 341  
 DB 880 tcatgctataatgcagacagatagatagctgtgacgtgacgtatlttgggagactaa 939  
 QY 342 cactcgatatacagaacttacc--aaatcagatgagtgatgagataagatgctct 398  
 DB 940 caccgactacagatgagatgctcctaaagataagaaatgagatgagatgagat 999  
 QY 399 gtgtttcaaaaacatgttacttcttcgattgctgctgctgctgctgctgctgct 458  
 DB 1000 gtcttttcagtaatgagacatttccacgtggtttatagtagattacacgtgttg 1059  
 QY 459 gaaactacccctagtgactatccacgtctcctctataatagtgcttcagcgagcatg 518  
 DB 1060 gaaactaccccttgatttgcacacagtcacgtcttcttgcacgttgcacgttgc 1119  
 QY 519 ttctagagatgctatcctatgacagtaaggaattagtccttaccatccacgtgggc 578  
 DB 1120 ggcgaagatatactcttacttactgataaagaactcttagcgataaagaacttgagc 1179  
 QY 579 tctgacagaagaagctgtgctatccatccgacagatcatalagcctttagggccagagaa 638  
 DB 1180 agtagctgaagaaggtcttgcgacgaattagacatgcatgttcattttaggagcaaaagaa 1239  
 QY 639 agaacttcaaatgatacacagaatccaagaatgcaagatttggcataaaaaggac 698  
 DB 1240 agaacttgaagaatgatacaaaaatcttagaagaagctaaagaatctgggataaagaagac 1299  
 QY 699 tatagctcaaaagtgtctctgtgtgctgtacttctttagatggaacttagact 758  
 DB 1300 tattacagcacaatatcttataagtgctgcttctcctgctgactatgacttattagctct 1359

QY 759 tgccttttgatggaactcctcttatttcttaatggagaactgtgatatcacatcgagac 818  
 DB 1360 ggcctcttggtatgagacacacttggctctctcagggga-----atattcattggaca 1413  
 QY 819 tttcttgctgtttcttctttagtgaatccatagcaagttatgattcatgtgagcagctcc 878  
 DB 1414 agaacacacgtatcttcttcttctgataatttgggctttagtggtagcagcatctcc 1473  
 QY 879 tcaattgaaacttcgcaataagcccgagagctgtccttcatatttccaggtattga 938  
 DB 1474 aagcattggaagcatttgcgaatgcaagagagagattatgaaatcttcaagataattga 1533  
 QY 939 taagaacccagatatagaatacttccacagctgtgatatcaacccgaatccatagaaag 998  
 DB 1534 taataagccaagatttgaacagcttctgaagtgaggccaacacgaataataatgaagg 1593  
 QY 999 aacttggaatttaaaaatgcttcttcaattatccatcaagaacacatcataagattct 1058  
 DB 1594 aaattggaattcagaagaatgcttacttccatccatccatccatccatccatccatcc 1653  
 QY 1059 gaaaggtctgaatcctagaatataagcttggagagacagctgccttggctgctcaatg 1118  
 DB 1654 gaagggcctgaacctgaagtgagagtgagtgagagtgagtgagtgagtgagtgagtgag 1713  
 QY 1119 cagtggaagagtgacgtgctgacgttctgcagaggttatatgacccgagatgagctt 1178  
 DB 1714 ctgtgggaagagacacacagctgacagctgagagagcttctatgacccacagagggat 1773  
 QY 1179 tatcatgtgtgagatgagatatacatcagagcttcaatgtgctgcatatccagacatat 1238  
 DB 1774 gttcagtggttgaatgagacagatatttggacataaatgtaagtttctcagggaaatcct 1833  
 QY 1239 tggagtggttgaatgagagagctgttcttggagacacacatcaagtaacaaatcagta 1298  
 DB 1834 tgggtgtgtgagtgacaggaacctgattgttggcaccacagatagctgaaacacttcgcta 1893  
 QY 1299 tggacagagatgagtgactatgaagaagtgagagagagagagagagagagagagagag 1358  
 DB 1894 tggcgttgaagaatgttcccatatgagatgagatgagagagagagagagagagagagag 1953  
 QY 1359 tgatttatcatgaggttctcctaataatttaatacatgttgtaggggaaagagagctca 1418  
 DB 1954 tgacttatcatgaaactgtcctataatttgaacaccttggtagggagagagagagagagag 2013  
 QY 1419 aatgagtgag 1478  
 DB 2014 gttagtggtgag 2073  
 QY 1479 gatctgattttagatgag 1538  
 DB 2074 gatcctcctgtgtgag 2133  
 QY 1539 agctgcactggaagaag 1555  
 DB 2134 ggtggtcctgagataag 2150

RESULT 9

AA294738  
 ID AA294738 standard; cDNA; 4646 BP.

XX AA294738;

DT 01-AUG-2000 (first entry)

XX Human ATP binding cassette ABCB1 (MDR1) cDNA.

XX ABCB1, ATP binding cassette; human; cholesterol; lipid disorder;

KW atherosclerosis; lipid disorder; dyslipidemia; porphyrinosis;

KW lupus erythematosus; diagnosis; gene therapy; MDR1;

KW multidrug resistance; chromosome 7q21; ss.

OS Homo sapiens.

```
XX      WO200018912-A2.
PN
XX      06-APR-2000.
PD
XX      21-SEP-1999;      99WO-EP06991.
PF
XX      25-SEP-1998;      98US-0101706.
PR
XX      (FARB ) BAYER AG.
PA
XX      Schmitz G, Klucken J;
PI
XX      WPI; 2000-293151/25.
DR
XX      Adenosine triphosphate binding proteins useful for identifying agents
PT      for treating atherosclerosis and other inflammatory disorders -
XX
XX      Claim 9; Page 110-112; 154pp; English.
PS
XX      The present sequence is that of human ATP binding cassette
CC      subfamily B protein ABCB1 cDNA. The cDNA was identified using a
CC      differential display method in which monocytes from peripheral
CC      blood were subjected to macrophage differentiation and cholesterol
CC      loading with acetylated low density lipoproteins and subsequent
CC      deloading with high density lipoprotein (HDL3) to identify
CC      cholesterol sensitive genes. The gene maps to chromosome 7q21
CC      and is also termed MDR1 (multidrug resistance). The invention
CC      provides cholesterol-sensitive ABC genes (see AA294734-63). These
CC      genes, and polypeptides encoded by them, can be used for diagnostic
CC      and therapeutic applications, and for biochemical or cell-based
CC      assays to screen for pharmacologically active modulator compounds
CC      useful for the treatment of lipid disorders, atherosclerosis or
CC      other inflammatory diseases such as psoriasis and lupus
CC      erythematosus.
CC
XX      Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other:
SQ
Query Match      31.1%; Score 603.8; DB 21; Length 4646;
Best Local Similarity 65.6%; Pred. No. 8.7e-171;
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;
```

```
Db      1180      agtagctgaagagctctgagcagcaattagactgtagtgcatttgaggacaaagaa 1239
Oy      639      agacttcaaaaggtatcacagaatctcaagaatgacaaggatttggcataaaagagac 698
Db      1240      agaacttgaaggtacacaaaatcttagaagaacttaagaatctggataaagaagac 1299
Oy      699      cataagctcaaaagtgctctctggtcgtgctgactctcttaagaaagaaactagact 758
Db      1300      tattacagcacaattattctataggtgctgcttctctgctgctatcatgacttaagcct 1359
Oy      759      tgccttttgatagtgaaacctcttgatctcttaatggaagaaacctgatatcacatcggaac 818
Db      1360      ggcctctgtagtggaaccacttgctctctcaccaggga-----atactattggaca 1413
Oy      819      tgtctctgcttcttcttcttagtgaatccatagcaagttatgcatctggagcagctccc 878
Db      1414      agtactcactgtatcttcttctgattatattggtggtctttagtggtagcagagcctccc 1473
Oy      879      tcaacttgaaaccttcgcaatagcccgagagctgcttcatatttccaggttatgta 938
Db      1474      aagcatgaaagcaatttgcacaatgcaagaagagcaagcttatgaaatcttcaagatatgta 1533
Oy      939      taagaacaaccagtatagataaacttccacagcttgatataaaccctgataatcaagaag 998
Db      1534      taataagccaagattatgacagctatctcgaaagatggtggcacaacaacagataatgaag 1593
Oy      999      aactgtggaatttaaaatgttcttcaattatccatcaagaacacatcatcaagattct 1058
Db      1594      aaatttgaatctcagaatagtgctcaacttgaatccacatctcgaaagaagttaagatct 1653
Oy      1059      gaaagcttgatattccagaatattgaatgagagacagctgcgcttgcttccaaagc 1118
Db      1654      gaaagcctgaaacctggaaggtgacagatggtggcagacggtggtccctggttgaaagatgg 1713
Oy      1119      cagtggaagagtaacgtatgctcagcttctgcagaagttatagatccagatgatgctct 1178
Db      1714      ctgtgggaagagcacaacacagctcagctgtagcaggtctatgacccacagagggagat 1773
Oy      1179      tatcatgttgatgataagatgacatcagaagctttaaattgctggcatctatcgagaccatat 1238
Db      1774      gttcagtgctgtagtgacaggaatatagagacaataatgtaaggttcttcaagggaaatcat 1833
Oy      1239      tggagtggttgtagcaaggccgtgttcttgtagggccacacacagtaacataataagta 1298
Db      1834      tgggtgtgtgtagtgcaggaacctgatactgtctgcccaccagaaagctgaaacatctgcta 1893
Oy      1299      tggacgagatgattgactgatatgaagagatgtagagagcaagcaagggaaagcaatgctta 1358
Db      1894      tggccgtgaaaaatgccaacatgtagatgtagaagatgtagaagctgccaaggaagccaatgcta 1953
Oy      1359      tgattttatcatbtaggttctcctaataatttaatacatgtgtaggggaaagagagctca 1418
Db      1954      tgactttatcatgaagaactgacctcaaaatttgaccacctggttggagagagaggggccca 2013
Oy      1419      aatggttgagggagcaagaagagagatggaatttctgtgctgaatttgaagaccacaa 1478
Db      2014      gttgagtggtgggagcaagcagaagatcgccaattgcaagtgccctgttgaacacccaa 2073
Oy      1479      gattctgattttagatgaggtacagctgcgacctgagattcagaagaagctcagctgttca 1538
Db      2074      gatcctcctgctgagtgagggccaagctcagcccttgtagcaacagaagcgaagcagtggttca 2133
Oy      1539      agctgcaactggaagaag 1555
Db      2134      gttggtctctgataaag 2150
RESULT 10
AAV32645
ID      AAV32645 standard; cDNA; 4669 BP.
XX
AC      AAV32645;
```



Db 1954 tgactttatcatgaactgcctcatcaatttgacaacctggttgagagaggagggccca 2013  
OY 1419 aatgagtgagggcagaacacagagatgcgaattgctgctcctagtctgaaacccca 1478  
Db 2014 gttgagtggtgagcagaagcagagatgcacatgacgtgctcctgctgcaacccca 2073  
OY 1479 gattctgatttagatgaggtacgtcgccttgatccagaagaagtaagctgttca 1538  
Db 2074 gatctcctgctgatatgagcagctcagccttgacacagaagaagcagatggttca 2133  
OY 1539 agctgcaactgagaaag 1555  
Db 2134 ggtgctctgataaag 2150  
RESULT 11  
AAT13394  
ID AAT13394 standard; DNA; 6505 BP.  
XX  
AC AAT13394;  
XX  
DT 24-JUN-1996 (first entry)  
XX  
DE Hybrid vector pSF-MDR.  
XX  
KM Hybrid; vector; gene transfer; gene therapy; haematopoietic stem cell;  
KM retroviral; murine embryonic stem cell virus; MESV;  
KM Mooney murine sarcoma virus; (MoMuSV);  
KM Friend murine leukemia virus; F-MuLV; ds.  
OS Synthetic.  
XX  
PN DE19503952-A1.  
XX  
PD 14-MAR-1996.  
XX  
PE 07-FEB-1995; 95DE-1003952.  
XX  
PR 08-SEP-1994; 94DE-1431973.  
XX  
PA (BOEF ) BOEHRINGER MANNHEIM GMBH.  
XX  
PI Baum C, Osterlag W, Stocking-harders C, Stockingharders C;  
XX  
DR WPI: 1996-152306/16.  
XX  
PT Hybrid retroviral vectors - for gene transfer into haematopoietic  
XX stem cells  
XX  
PS Disclosure: Page 25-29; 42pp; German.  
XX  
CC New hybrid vectors comprise (1) a leader region including the U5  
CC region and tkRNA primer binding site of murine embryonic stem cell  
CC virus (MESV) or Mooney murine sarcoma virus (MoMuSV), and (2) a 3'-  
CC LTR including the U3 and R regions of a Friend murine leukemia  
CC virus (F-MuLV). The vectors are useful for ex-vivo or in-vivo gene  
CC therapy. High levels of gene transfer can be achieved in  
CC haematopoietic stem cells and their myeloid (non-lymphatic) progeny.  
CC pSF1, pSF2, pSF3 and pM1 (sequences given in AAT13390-T13393) are  
CC examples of such vectors.  
CC Vector pSF-MDR (sequence given in AAT13394) is based on the  
CC MESV vector R224.  
XX  
SQ Sequence 6505 BP; 1719 A; 1506 C; 1606 G; 1674 T; 0 other;  
Query Match 31.1%; Score 603.8; DB 17; Length 6505;  
Best Local Similarity 65.6%; Pred. No. 1.1e-170;  
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;

OY 222 gattcctgtggaattataactgcagccagcagacccaagagatctcgaaacagtttt 281  
Db 2212 gtttcaatttgggtccggtgagcgtggaagacaataacacaaattagaacagtttt 2271  
OY 282 tcaatcagttttgacacagacatcgcgtgttgatagctgtgtacatcggtgaactaa 341  
Db 2272 tcatgtctaattgagacagcgagatagctgtgttgatgtgacagatgtgtggagcttaa 2331  
OY 342 cacttcgcatcagacatctgac---aaacatcagatgtgtatgtggaataagattgctct 398  
Db 2332 caccgacttaacagatgagtctcctaagattaaatgaattatgtgtgacaataatgtgaat 2391  
OY 399 gttgtttcaaacatgtctacttttgcattggcctcgagcttggtgtgtggaagggctg 458  
Db 2392 gttcttcaagcaatggaacatcttccacggtgttataagatgaattacaagttgttg 2451  
OY 459 gaaactcacctagtgatctatccacgctcctccttataatgtgttcacgagcagcatg 518  
Db 2452 gaagtaacctgtgattgttgccatcagtcctgttcttgagctgtcagctgtgtctg 2511  
OY 519 ttctaggtgtcatctcatctgacagtaaggaatgaatgtcctattccaaagctgggac 578  
Db 2512 ggcacaagatactatcttcaattactgataaagaactcttaagatgataaagctggagc 2571  
OY 579 tgtgccaagaagaagctgtgcatcaatccagacagctatagcctttagggcccaaggaa 638  
Db 2572 agtaagcgaagaagctgtgagcagcaatlaagacatgtgattgattggaacaaagaa 2631  
OY 639 agaactcaagaagatacacagaactcacaagaatgcagaagatttggcataaaagagac 698  
Db 2632 agaacttgaagaagatcacacaaataatttagaagaagctaaagaatttggataagaagac 2691  
OY 699 tatagcttcaaaagtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 758  
Db 2692 tatcacagccaatttctatagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2751  
OY 759 tgccttttgataggaacctgtgtatcttcaattgaatgaagacctgtgatatacattgggac 818  
Db 2752 ggccttcgtgattggaccacctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2805  
OY 819 tgtctgt 878  
Db 2806 agtaactcgt 2865  
OY 879 tcaattgaaccttcgaataagcccgagagagctgtccttcaatattccaggttatgga 938  
Db 2866 aagcattgaagcatttgcacaatgcagaagagagcagcattatgaatcttcaagataatga 2925  
OY 939 taagaaccccgatagataaactttccacagctgtgatatgaacctgtatccatagaag 998  
Db 2926 taataagccaagatattgacagctatccagaagagtgaggcaacacagataatataag 2985  
OY 999 aacttggaatttaaaatgttcttctcaatcatatccatcaagaacatcatcaagattct 1058  
Db 2986 aaatttgaattcagaataatgttcttcaatcatatccatccatccgaagaagttaaatcct 3045  
OY 1059 gaaaggttgaattcagaataatgttcttcaatcatatccatccatccgaagaagttaaatc 1118  
Db 3046 gaagggcctgaacctgaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 3105  
OY 1119 cagtggaagagtagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1178  
Db 3106 ctgttggaagagtagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 3165  
OY 1179 tatcatgtgtgagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1238  
Db 3166 gttcgt 3225  
OY 1239 tggagtggttagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1298  
Db 3226 tgggt 3285





Accession	Sequence	Length
Db 2623	ttgtgtgtgtgaagtcagaacctgtattgtttccaccacgaagctgtgaaacattcgcta	2688
QY 1299	tgagcgaatgatgtgtcattgtatgaagagatlgagagagcagaagaggagacaattcgta	1358
Db 2689	tgagcgtggaatatgtccaccatgatgatgtaggaaagtgtcgaaggagccaatgcta	2748
QY 1359	tgattttatcatgtgaagtttccataataattaatcatattgtgtaggagaaaaggagctca	1418
Db 2749	tgactttatcatctgaacattgcctcataaatttgaacacctggtcttgagagagagggccca	2808
QY 1419	aattagtggaagggcagaacaagagatgcgaattgtctgtcttgacttaatttcgaaccccaa	1478
Db 2809	gttgaagtggtgggcagaagaagagatgcgcatttgcacgttgccctgglttcgaaccccaa	2868
QY 1479	gattctgatttagatgaagctacgctctgcctctgagattcaagaagaagtcagctgttca	1538
Db 2869	gattcctctcgtgatgagggcagcagtcagccttgagacacagaagaagcgaagcagtgttca	2928
QY 1539	agctgcactctggagaag 1555	
Db 2929	ggttgctcttgataag 2945	

RESULT	1.3
AAZ49333	
ID	AAZ49333 standard; cDNA; 3860 BP.

AC	AAZ49333;
XX	
DT	14-MAR-2000 (first entry)

DE Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.

KW Multidrug resistance; MDR-1; P-glycoprotein;  
 KW transmembrane efflux pump; haematopoietic stem cell; transduction;  
 KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;  
 KW gene therapy; gene replacement; genetic defect; thalassemia;  
 KW Gaucher's disease; sickle cell anaemia; leukæmia; ex vivo expansion;  
 KW cytokine; mutant; ds.

OS	Synthetic.
OS	Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	1..3843

FT /product= "Human G185V mutant MDR-1 protein"

PN W09961589-A2.

PD 02-DEC-1999.

PF 27-MAY-1999; 99WO-US11825.

PR 28-MAY-1998; 98US-0086988.

PA (SUUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Sorrentino B, Bunting K;

DR WPI; 2000-072615/06.

DR P-PSDB; AAY58187.

PT Ex vivo expansion of hematopoietic stem cells transduced with a  
PT sequence encoding human multidrug resistance-1, used for bone marrow  
PT transplantation -

PS Example 1; Page 79-82; 113pp; English.

CC This sequence represents cDNA encoding human G185V mutant multidrug  
CC resistance protein MDR-1, where the Gly residue at position 185  
CC of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a

transmembrane efflux pump, responsible for the export of drugs from cells, particularly cancer cells. The wild-type MDR-1 shows increased resistance to etoposide and decreased resistance to vinca alkaloids compared with the G185V mutant. The invention relates to transducing haematopoietic stem cells with nucleic acid encoding an MDR protein and culturing the modified cells. The modified haematopoietic stem cells are useful in bone marrow transplantation (to reconstitute haematopoietic systems in patients who have undergone chemotherapy or radiation therapy) and in ex vivo gene therapy of genetic defects in cells derived from haematopoietic stem cells, e.g., thalassemia, Gaucher's disease, sickle cell anaemia or leukaemia. The modified cells can also be used to identify factors involved in regulating proliferation and differentiation in haematopoietic stem cells. Haematopoietic stem cells that express MDR-1 will be protected against chemotherapeutic agents, so can be engrafted while the patient is undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells provides sufficient cells to permit standard biochemical analysis. Overexpression of MDR-1 allows cytokine-driven expansion of haematopoietic stem cells by at least 10-fold compared with a maximum of 4-fold in known procedures.

Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Query Match	31.0%	Score 602.2;	DB 21;	Length 3860;
Best Local Similarity	65.5%;	Pred. No. 2.4e-170;		
Matches 915; Conservative	0;	Mismatches 473;	Indels 9;	Gaps 2;

QY	162	caggttggcccgctgattatgtttggaatagtggttcgcgccttgaattttgtttgatacaca	221
Db	336	cagataagccataattatacagtggaatttgggtgcgggggtgcgcgtgcgtccttaccatacca	395
QY	222	gatttcctcttgtagataatactgacagccagcagcaaccaagaagaggtttcgaaacaagttttc	281
Db	396	ggtttcattttgtgtgcctgcgcgcgtgcggagcaaatacacaacaattagaagaacgtrttt	455
QY	282	tcattcagttttggcacagagacatcgcgtgtgtttgatacgtctgacatcgtgtgaacttaa	341
Db	456	tcattcctaatagcacacagagagataagctctgtttgatatgtgacagatgttggagagcttaa	515
QY	342	cactgcagatcacagacatttgcg----aaatacagtagatgtatgtgaatagaattgtctc	398
Db	516	caaccgcgaacttaccagatgtgtcctctaagataataagaaattatattgtgtgacaaatttgaat	575
QY	399	gtttgttcaaaaacatgtctacttttctcgatgtgcctgcggcagttgtttgttgaaaggtcg	458
Db	576	gtctttcttaagtaacatgycacaatttttttcaactgggttttaataagatttacaacgtgttg	635
QY	459	gaaacttcagccctagtagatctatccatccagcttcgcctcttaatatgtcttcaacggcagatg	518
Db	636	gaagctaaccccttggattcttggccaatcagctccgttcttccttgagctgcacgcctgcctcg	695
QY	519	ttctcgtatgtgcatactcatctatgcacagtaagaagaaattaaagtccattccacaagcttgggc	578
Db	696	ggcgaagatactatcttcatcttacttaccagataaagaacctcttagcgtatgcaaaagctgtgagc	755
QY	579	tgttgcagagaagaagctcttgcataccaatccgcgaacagctcaatagctttagggcccaaggaa	638
Db	756	agtagctgaaggggtctcttgcgcagacataagaaacgttgattcattcttggaaagacaaga	815
QY	639	agaacttcaaaagttatcacaggaactctcaagaatgtgcagaagtaatttggctataaaagac	698
Db	816	ggaactctgaagaagttacaacaaaattttagaagaagctcaaaagaatttggatataagaagc	875
QY	699	tataagcttcaaaagtgctctctctgtgtcgtgtgttaactctttagatgtgaagaaactatgact	758
Db	876	tattacagcaaatatttctaatagtgtgcgtcttccgtctgactatgacatcttatgctc	935
QY	759	tgtcttttgggtatgtgaacctccttgattcttcaatgaggaacctgtgatataccatcgggac	818
Db	936	ggcctctctcgtatattggagcaacctgtgtctctccacgggga-----atattctatggaca	999
QY	819	ttctcttcctgtgtttcttcttaagtaatccacagagaagttatttgcatttgaacagacgtccc	878

```

Db      990 agtaccacacgtatcttctctgtatatttggtggtttagtggtgacagcatctcc 1049
QY      879 tcaattgaaaccttcgcaataagccgagagctgcttccattatcttcagggttatga 938
Db      1050 aagcatggaagcatttgcacaaatgagagcagcttaagaaattctcaagaataatga 1109
QY      939 taagaaccacagatataatactttccacagctggtatataaactgaatccatagaag 998
Db      1110 taataagccaaagtatttgacagctatccgaagatggtgcaacaaccagataataagg 1169
QY      999 aactgtggaatttaaaatgttcttccaattatccacccaacatctcaagaattct 1058
Db      1170 aaattggaattcagaagaatttccactcagttacccatcccgaaagaagtaagatctt 1229
QY      1059 gaagagctcgaatctcgaatgaatgctgagagacagctgcctgtgtgtctcaatgg 1118
Db      1230 gaagggccttgacactggaagtgacagtggtgacagcgtgtgctgtgtggaacagtg 1289
QY      1119 cagtgggaagatagcgttagtccagctctgcagaggttatatgacccgagtatgctt 1178
Db      1290 ctgtgggaagaagcacaacacatccagctgacgagagctctatgacccacagaagagat 1349
QY      1179 tatcatgtgtgagatgagatgacatcagagctttaaatgtgcggtatctcgaagacat 1238
Db      1350 gctcagctgtgagacagagatatgaaccataaattgaagttctacgggaatcat 1409
QY      1239 tggagtggttagcacaagagcgtgtgtgttcgagacacacatccatgataatcaagta 1298
Db      1410 tgggtgtgtgagtcagaagaccgttatgttgcacacacagatagctgtaaacatctgcta 1469
QY      1299 tgaagcagatgattgtgactgataagaagatgagagagcagaagggaaacaaatgcgta 1358
Db      1470 tggccgtggaagaattgtccacatgtagatgtgagaagcgtgtcaaggaagcaaatgtcta 1529
QY      1359 tgattttatcatgagatcttccctaataattatatacatgtgagggaaaaagagctca 1418
Db      1530 tgactttatcatgaactgctctcaataatttgacacacctgtgtgagagagagggccca 1589
QY      1419 aatgaatggagggcagaagaagaagatcgcaatgtcgtgccttagttcgaaaccccaa 1478
Db      1590 gtctgagctgtgagcagaagcagaagatcgccatgcagctgtgcctgtcgaaccccaa 1649
QY      1479 gatctgattttagatgagctacgtctgcctcgtgattcagaagaagcagtcgttca 1538
Db      1650 gatccctcgtcgtgattgagggcagctcagccttggaacacagaagaagcagagtggtta 1709
QY      1539 agctgcactggagaag 1555
Db      1710 ggtgctctgataag 1726

RESULT 14
AAN70752
ID      AAN70752 standard; cDNA; 4669 BP.
XX
AC      AAN70752:
XX
DT      21-MAY-1991 (first entry)
XX
DE      Sequence of human multi-drug resistance-1 (mdr1) cDNA from clones
XX      lambda-HDR10,5 and 104.
XX
KM      Chemo-therapy resistant tumour cell; P-glycoprotein; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS 425..4267
XX      /tag= a
XX      PN      WO8705943-A.

```

```

PD      08-OCT-1987.
XX
XX      26-MAR-1987; 87MO-US00758.
PF
XX      01-AUG-1986; 86US-0892575.
PR
XX      28-MAR-1986; 86US-0843610.
PA
XX      (UNIL ) UNIV OF ILLINOIS.
PI
XX      Roninson IB, Pastan IH, Gottesman MM;
DR      WPI; 1987-291656/41.
DR      P-PSDB; AAP70452.
XX
PT      DNA for multi-drug resistance in human cells - used to detect
PT      chemotherapy-resistant tumour cells and for producing
PT      polypeptide(s) for diagnosis and therapy
XX
PS      Claim 2(a); Table 5, pp30-39; 61pp; English.
XX
CC      The human multi-drug resistant KB carcinoma cell lines were used as
CC      the source of the mdr1 gene nucleic acid sequences (AAN70751). To
CC      obtain cDNA clones of the mdr1 gene (AAN70752), poly (A) and RNA was
CC      used. Analysis of the AA sequence presented in (AAN70752) indicates
CC      that the mdr1 gene product is likely to be a transmembrane protein.
CC      The presence of transmembrane domains and potential glycosylation
CC      sites is consistent with the mdr1 protein being related to the
CC      P-glycoprotein.
XX
SQ      Sequence 4669 BP; 1393 A; 896 C; 1126 G; 1254 T; 0 other;

Query Match      31.0%; Score 602.2; DB 8; Length 4669;
Best Local Similarity 65.5%; Pred. No. 2,7e-170;
Matches 915; Conservative 0; Mismatches 473; Indels 9; Gaps 2;

QY      162 caagttacccctgattatgttggaaatagtgctgctcgttatttctgttaacata 221
Db      760 caggtatgcctattatcatcagatggaattgctgctggtgtgtgtcttacttaca 819
QY      222 gatttctctgtgatttaactgacagacagacacaaagagatcggaaaacagtttt 281
Db      820 gtttcatlcttgtgctgtgcagctggaagacaataacaaattagaacaacgctttc 879
QY      282 tcaatcagtttggcacaagacatcgctgtttgataagctgtacacgcgtgaacttaa 341
Db      880 tcatgtcataatgycagagagatagagctgtttgagtgcaagatgttggagagctaa 939
QY      342 cactgcatagcacagacattgac--aaatcagtgatgtattggaataagatgtcct 398
Db      940 caccgacttacagatgatgtctctaagattcaagattatgtgtgtaacaaattggaat 999
QY      399 gttgtttcaaaacatgctactcttttcgattgctgcctgagtggttggtaaggctg 458
Db      1000 gttctttcagtcacatgycacacatttccactgggtttatagatggaattcacgtgtg 1059
QY      459 gaaactcacccatgtagctatccacacgctctccctcatataatgagtcagcggacgat 518
Db      1060 gaagctaacctctgtgtatttggcatalcagactcgttcttgagctgcaactgtcgtctg 1119
QY      519 ttctagatgtgcatcattgacagtaagaaatagtgctatctccaagctgggc 578
Db      1120 gycaaagatactatcttcatcttactgataagaactcttagcgaatgcaaaagtggagc 1179
QY      579 tgtgcaagaagaattctgtcatcaatccgaacagctatgctttagggcccgagaa 638
Db      1180 agtagctggaagagctctgycagcaatlaagaactgtgattgcaattggaagcaagaag 1239
QY      639 agaaactcaaggtatcacagaatcctcaagatgacgaagatttttgccataaaagac 698
Db      1240 agaaactggaaggtacacaaaatttagaagaagcctaaagaatttggataagaagac 1299
QY      699 tatagctcaaaagtgtctctgtgtgtgttacttcttataagaaatgaacctaagact 758

```

```

Db 1300 tattaacagcaaatattctatagagtgctgcttctctgcatcatalgtatctatgctct 1359
Qy 759 tgccttttgatagtgaaacctcttgatcttcttaatgagaacctgatatatacatcgagac 818
Db 1360 ggccttctgatatgagcaacctgtgtctctcagggga-----atatctattgagca 1413
Qy 819 tcttctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 878
Db 1414 agtaactcattatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 1473
Qy 879 tcaattgaaaccttcgaataagcccgagagctgcttccattacatttccaggttatgta 938
Db 1474 aagcattgaagcatttgaaatgcaagagcgagcagcttatgaaaccttcaagatatgta 1533
Qy 939 taagaacaccagtatagataacttccacagctgagatataaactgataccatgaagaag 998
Db 1534 taataagccaagtattgacagctattcgaagaagtgagcacaacacgaataataagaag 1593
Qy 999 aactgtgaatttaaaaaagtcttcttcttcttcttcttcttcttcttcttcttcttctt 1058
Db 1594 aaatttgaattcagaataatgctcacttcaagttaaccatctcgaataagaagttaagatct 1653
Qy 1059 gaaaggtcgaatcgcagaataatgctcgcagagacagctgctgctgctgctgctgctgct 1118
Db 1654 gaagggcttgaacctgaaagtgagagtgagcgagcgagcggtgctgctgctgctgctgctg 1713
Qy 1119 cagtgaggaagtgacggtagctcagctctcgcagagtgatataatgacggatgagctt 1178
Db 1714 ctgtgaggaagtgacaaagtgacagctgagtgagcgctctatgacccacagagggagat 1773
Qy 1179 tatctgtgtgagtgagtgatgacatgacagctttaaattgagcgcttcttctgagaccat 1238
Db 1774 ggtcagtgctgagtgagcagagatataagacataatgtaagtgcttcttcttcttctt 1833
Qy 1239 tggagtgctgagtgacagagcgtgttcttgcgagacccacacgaatacaaatcaagta 1298
Db 1834 tgggtgtgtgagtgacagagcgtgtatgttgcacacgaatagctgtaaaatcctgcta 1893
Qy 1299 tggacgagatgagtgtgacatgagtgagagtgagagcagcagaaggaagcaaatgctgta 1358
Db 1894 tggcgtgtaaaatgacacatgagatgagtgagtgagtgagtgagtgagtgagtgagtgag 1953
Qy 1359 tgatttctatgagtggttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1418
Db 1954 tgaattctatcgtgaacgtgctcatalaattgacacccctggttgagagagagggggccca 2013
Qy 1419 aatggtgagagcgagaaacagagatcgcaattgctgctgcttgaattcgaaaccccaa 1478
Db 2014 gttgagtggttgagcgagagcagagagatcgccatgacagtgctgctgctgctgctgctg 2073
Qy 1479 gatcttgatttagatgagtgagtgctgctgctgctgctgctgctgctgctgctgctgct 1538
Db 2074 gatctctctgctgagtgagcgacagctcagccttgagacacagaagaagcgagcgtgttca 2133
Qy 1539 agctgactgagagaag 1555
Db 2134 ggtggtcttgatagaag 2150

```

## RESULT 15

AA052726 standard; cDNA: 4669 BP.

AA052726;

24-JUN-1994 (first entry)

Sequence of human multi-drug resistant gene mdrl.

p-glycoprotein; multi-drug resistant gene; mdrl; ss.

Homo sapiens.

```

XX Key Location/Qualifiers
FH 424..4267
FT CDS /*tag= a
PN WO9324613-A.
PD 09-DEC-1993.
PF 14-MAY-1993; 93WO-US04707.
PR 22-MAY-1992; 92US-0887712.
PX (GENE-) GENETIC THERAPY INC.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI McDonagh KT, Nienhuis A, Tolstoshev P;
XX WPI: 1993-405805/50.
DR P-PSDB: AAR44297.
XX DNA or RNA sequence for human multi-drug resistant gene MDRL -
PT encodes P-glycoprotein with altered splice site, used in cancer
PT therapy
XX Example: Fig 4; 64pp; English.
XX PMDR2000 contains an mdrl cDNA sequence (AA052726) described in PCR
CC application no. WO87/0943m wherein the first 282 bp of the 5' UTR
CC and the last 23 bp of the 3' UTR of the cDNA sequence have been
CC removed. The codon YTN is unclear in the spec., and encodes the AA
CC leu.
SQ Sequence 4669 BP: 1394 A; 893 C; 1126 G; 1254 T; 2 other:

Query Match 31.0%; Score 602.2; DB 14; Length 4669;
Best Local Similarity 65.5%; Pred. No. 2.7e-170;
Matches 915; Conservative 0; Mismatches 473; Indels 9; Gaps 2;

Qy 162 caggttgaccctgtatattgttggaatagtggtgctgctgctgcttattgtgtatcata 221
Db 760 cagatgacatctattatcaagtgagatgtggtgctgctgctgctgctgctgctgctgctgct 819
Qy 222 gattccttctgtgattataactgacgacgacgacgacgacgacgacgacgacgacgacgac 281
Db 820 gtttcatcttctgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 879
Qy 282 tcatcagtttgacagagagatcgtgctgctgctgctgctgctgctgctgctgctgctgctgct 341
Db 880 tcatgctataatgacagagagatagctgctgctgctgctgctgctgctgctgctgctgctgct 939
Qy 342 cactgcgcatgacagatgac---aaatcagtgatgtatgagataagatgctct 398
Db 940 caccggacttaagatgattgtctcctaagataatgaacttaattgttgcacaaattgagat 999
Qy 399 gttgttcaaaaacatgctacttcttctgattgctgctgctgctgctgctgctgctgctgctg 458
Db 1000 gttcttcaagtaatgcaacatttcttcaattgcttgaattagatgagatgacgtgtgctg 1059
Qy 459 gaaactcaccctagtgactctatccagctctcttataatgagcttgaagcgacgctg 518
Db 1060 gaagtaaaccttgattgttgccataagctcgttcttctgagctgacgtgctgctgctg 1119
Qy 519 ttctagatgctacatcatttaccagtaagaatgaagtgtcctatccaaagctgggc 578
Db 1120 ggcacaagatcattcttcaattcagatgaagaactcttagtgatgcaaaagcgtgggc 1179
Qy 579 tgtgcagaagaagctctgtgcatcaatccgaacagatcagcctttagggccagagaa 638
Db 1180 agtaagtgaagaggtctgacgacatgaactgtgattgtattgagagcaaaagaa 1239
Qy 639 agaattcaaaagtatacacaagatctcaaaagtacaaagatttggcatataaaggac 698

```

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:32:35 ; Search time 9392.76 Seconds  
(without alignments)  
2219.455 Million cell updates/sec

Title: US-09-873-409-16

Perfect score: 1940  
Sequence: 1 ttccgcttgcgcgatgact.....taaatgcctgacttatgga 1940

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpl:\*  
6: em\_estlba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	545	28.1	545	10	AL040762
2	489.8	25.2	533	12	N24315
3	489.8	25.2	533	12	N24315
4	482.8	24.9	559	11	BF692596
5	385.8	19.9	405	10	AA243820
6	378.8	19.5	872	12	AK020318
7	296.2	15.3	894	11	BF584668
8	253.8	13.1	803	11	BG298756
9	233.6	12.0	780	11	BG587938
10	221.8	11.4	943	10	AL520322
11	219.8	11.3	661	11	BG302675
12	219.4	11.3	932	13	AZ670821

C 13	218.6	11.3	880	13	AZ687805
C 14	213.8	11.0	575	10	AM342644
C 15	209.2	10.8	886	13	AZ540627
C 16	209	10.8	853	13	AZ679807
C 17	200.8	10.4	891	13	AZ682250
C 18	199.8	10.3	787	11	BG584063
C 19	197.8	10.2	402	11	H82550
C 20	196	10.1	886	13	BH139685
C 21	191.2	9.9	871	13	AZ682350
C 22	189	9.7	947	13	AZ683753
C 23	188	9.7	1019	11	BG248052
C 24	187.8	9.7	885	13	AZ686798
C 25	187.6	9.7	860	13	AZ543293
C 26	185	9.5	750	11	BG585786
C 27	184.8	9.5	611	11	BF635924
C 28	184.6	9.5	687	11	BG646725
C 29	184.4	9.5	834	13	AZ548312
C 30	183.8	9.5	503	13	AQ380214
C 31	183.6	9.5	897	13	AZ541090
C 32	181.2	9.3	656	10	AM686102
C 33	178.8	9.2	823	13	AZ532602
C 34	177.8	9.2	878	13	AZ539524
C 35	176.2	9.1	651	10	BE016688
C 36	172.4	8.9	921	13	AZ687628
C 37	167.6	8.6	795	11	BF313560
C 38	167.4	8.6	919	13	AZ690701
C 39	166.8	8.6	944	11	BF796582
C 40	166.2	8.6	785	10	AV709991
C 41	164.4	8.5	815	11	B1409057
C 42	164.4	8.5	2676	12	AK014319
C 43	163.4	8.4	842	13	AZ671925
C 44	161.2	8.3	541	10	A1722858
C 45	160.6	8.3	899	13	AZ548799

#### ALIGNMENTS

RESULT 1  
AL040762  
LOCUS  
DEFINITION DKEFp434C1815\_r1 434 (synonym: htes3) Homo sapiens cDNA clone  
DKEFp434C1815 5', mRNA sequence.  
ACCESSION AL040762  
VERSION AL040762.1 GI:5409708  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 545)  
Bloecker, H., Boecher, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and  
Wiemann, S.  
EST (Bloecker, et al.)  
JOURNAL Unpublished (1999)  
COMMENT  
Contact: Bloecker H

MIPS  
Am Kioferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKEFp434C1815) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..545  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKEFp434C1815"

#### FEATURES

source





```

OY 1894 tctaagaaaatcgagcgtcttttttaaatgctgacttatga 1940
Db 481 CCTAAGAAAATCCAGGCTTC-TTTTAAATGCTGGCTTATGGA 526

RESULT 3
LOCUS U66692 533 bp mRNA HTC 23-JUL-2001
DEFINITION Homo sapiens clone EST422562 mRNA sequence.
ACCESSION U66692
VERSION U66692.1 GI:1906577
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Allikmets, R., Gerrard, B., Hutchinson, A. and Dean, M.
TITLE Characterization of the human ABC superfamily: isolation and
JOURNAL mapping of 21 new genes using the expressed sequence tags database
MEDLINE Hum. Mol. Genet. 5 (10), 1649-1655 (1996)
97049974
REFERENCE
AUTHORS Allikmets, R., Gerrard, B. and Dean, M.
TITLE Direct Submission
JOURNAL Submitted (12-AUG-1996) Human Genetics Section, National Cancer
Institute, NCI-FCIDC, Frederick, MD 21702, USA
FEATURES
source
1. 533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="EST422562"
/Note="similar to ATP-binding cassette transporter"

BASE COUNT 139 a 134 c 118 g 142 t
ORIGIN

Query Match 25.2%; Score 489.8; DB 12; Length 533;
Best Local Similarity 98.1%; Pred. No. 5.3e-123;
Matches 517; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

OY 1416 tcaatgagatgagggcagaagaagagatcgcaattgctgctgacttagtgcgaacc 1475
Db 1 TCATAAGATGAGGGAGGCGCAGAAACAGAGATCGCAATTCGCTGCTTACGTTGGAACCC 60

OY 1476 caagattctgatttagatgaggtactgctgctgctgactgacaaagcaagtcagctgt 1535
Db 61 CAAGATTCTGATTTTATGATGAGGCTACGCTGCTGCTGATTCAGAAAGCAAGTCAGCTGT 120

OY 1536 tcaagtgcactgagaaagagatcccccaaggtatcattttgacctaatcaccacag 1595
Db 121 CTGAGCTGCATGGAAGAAGATACCCCAAGTATTCATTTTGAACCTAACCTCAAG 180

OY 1596 tggagatcgctgacttgaaccagcgcccttgacagactctgccccctcaaacccacc 1655
Db 181 TGGAGAAATCGCTACCTTGAAACGAGCGCCCTTGACAGCTCTGCCCCCTTAACCTCAC 240

OY 1656 ctgacctctgctgactatgagctactgacacatcctcaagcgcatatgacgttggcc 1715
Db 241 CTGACCTCGCTGCTGCTATGAGCTACTGACACATACCTCAAGGCGCATATGCAAGTTGGCC 300

OY 1716 ctgacccaacttaactgaatcagaagggaggtgagcagtgcgtagaagaaacacat 1775
Db 301 CTGACCAAAATTAACCTGATCTAGAGGGGAGTGGCACTGGCGGTATGAAAAAACCAT 360

OY 1776 gaacagtttctcgatgagcctgactccctataaaccagagccttgacccttcaaa- 1834
Db 361 GAACAGTTTCTCGATGAGCTGACTCCCTTATTAACCAAGGCTTACAGACCCCTTACAAA 420

OY 1835 -ggcttaatgacacatttacttctgacttgccttggaagtgaagcgttttttttc 1893
Db 421 GGCTTAATGAGGACATTTTACTTGCATTTGCTTGAAGTGAAGTAAAGCGTTTTTTTC 480

```

```

OY 1894 tctaagaaaatcgagcgtcttttttaaatgctgacttatga 1940
Db 481 CCTAAGAAAATCCAGGCTTC-TTTTAAATGCTGGCTTATGGA 526

RESULT 4
LOCUS BF692596 559 bp mRNA EST 22-DEC-2000
DEFINITION BF692596 602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
ACCESSION BF692596
VERSION BF692596.1 GI:11978004
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgi.nhl.nih.gov/
TITLE 1 (bases 1 to 559)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCRD/DMP
CDNA Library Preparation: CLOVENTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM1203 row: a column: 13
High quality sequence stop: 553.
FEATURES
source
1. 559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4334100"
/clone_11b="NIH_MGC_62"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (p1 phage-resistant)"
/Note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgccctggcc); Site_2: SfiI (ggcattatggc);
Double-stranded cDNA was prepared from cell line RN. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGAGCGCGCCGACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

```

BASE COUNT 161 a 118 c 141 g 139 t
ORIGIN

Query Match 24.9%; Score 482.8; DB 11; Length 559;
Best Local Similarity 96.8%; Pred. No. 4.4e-121;
Matches 546; Conservative 0; Mismatches 12; Indels 6; Gaps 5;

OY 1159 tatgacgcgagatgacttatcatggtgagagagatgacatgagactttaaagt 1218
Db 1 TATGATCCGGATATGAGCTTATACCGTGTGATGAGATACATCAGCTTTAAATG 60

OY 1219 cggcatlatacagacatattgagtgagtgagtcgaagcgttctgttcggaccac 1278
Db 61 CGGCATTATGAGACATTA-TGAGTGTGTTAGTCAAGAGCTG-TTTGCTCGGAGACACC 118

OY 1279 atcaagtaacaatcgaagtagcagagatgacgtgacgtgagaaagatgagagca 1338
Db 119 ATCAGTAACAATATCAAGTATGAGACGAGATGATGACTGATGAAGATGAGAGAGCA 178

OY 1339 gcaagggaagcaatgctgtagtttatcatgaggttccataaattaatcaatg 1398

```

DB	179	GCAGGAAAGCAATATGATATGATTTTATTCATGGAGTTTCTCTAATTAATTAATATGATTC	238
OY	1399	gtagggggaaaa-aggagctcaaatgagtgaggggagcaagaagagatcgcgaatgctgcg	1457
DB	239	GTAGGGGAAAAACAGAGAGCTCAAAATGAGTGGAGGGCAGAAACAGAGATTCGCAATTCGTCG	298
OY	1458	tgccctaatctgaaaccccccaagatctcgatcttagatgaagaagctgaagctgcgcctgagatc	1517
DB	299	TGCCCTTAAGTGGAAACCCCAAGATTCATATTTTACATGAGGCTACGTCGTGCCCTGGATTC	358
OY	1518	agaagaagcaagctgaagctgctcaagctgcgaactgagagaagatcccccaagatctcatctg	1577
DB	359	AGAAAAGCAGATCAGCTGTTCACAGCTGCACCTGAGAGAGATACCCCCAGATTCATTTTTCG	418
OY	1578	acctaatctcaactcaagtgagaaatgcgctgcaccttgaaacagcgccctcgacagctct	1637
DB	419	ACCTTAATTTACCTCAATGAGATGGAGATTCGCTGACCTTGAACAGAGCCCTTCGACAGCTCT	478
OY	1638	ggccctcaaacctcaactcaactgcctccctgcctcgtcctatgagctactgcacatacctccaag	1697
DB	479	GGCCCCCAAC--TCACCTGACCTCCGCTGACATGATGAGCTACAGCACAATACCTCAAGG	536
OY	1698	ccatagcatgtgctgcctcgac	1721
DB	537	-CATATGCAGTTTGTGGCTGAC	559
RESULT	5		
AA243820			
LOCUS	AA243820	405 bp	mRNA
DEFINITION	zr7g9g6.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:668506 5'	EST	06-AUG-1997
	similar to SW:MDRI_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ; ,		
ACCESSION	AA243820		
VERSION	AA243820.1	GI:1874631	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 405)		
	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,		
	Kucaba,T., Lacey,M., Le,N., Lennon,G., Maris,M., Martin,T., Moore,B.,		
	,T., Waterston,R. and Wilson,R.		
	WashU-Merck EST Project 1997		
	Unpublished (1997)		
TITLE	Contact: Wilson RK		
JOURNAL	Washington University School of Medicine		
COMMENT	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	This clone is available royalty-free through LNL; contact the		
	IMAGE Consortium (info@image.lnl.gov) for further information.		
	Insert Length: 1572 Std Error: 0.00		
	Seq primer: -28m13 rev2 ET from Amersham.		
FEATURES	Location/Qualifiers		
source	1..405		
	/organism="Homo sapiens"		
	/db_xref="GDB:5562473"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:668506"		
	/clone_1ib="Soares_NHMPu_S1"		
	/tissue_type="Pooled human melanocyte, fetal heart, and		
	pregnant uterus"		
	/lab_host="DH10B"		
	/note="Organ: mixed (see below); Vector: pT7T3-Pac		
	(Pharmacia) with a modified polylinker; Site 1: Not I;		
	Site 2: Eco RI; Equal amounts of plasmid DNA from three		
	normalized libraries (melanocyte 2NbH, pregnant uterus		
	NbHPU, and fetal heart NBH19W) were mixed, and ss circles		
	were made in vitro following HAP purification, this DNA		

was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.C.E. clones 262232-265223, 340488-345479, and 484488-489479."

Query Match	Similarity	19.9%	Score 385.8	DB 10	Length 405
Best Local	Similarity	99.5%	Pred. NO. 1.4e-94		
Matches 387	Conservative	0	Mismatches 2	Indels	Gaps 0
Oy	822	ctcttcgtgttcttcttaagtatccatcagaagcttgcattgcattgagcagcagctccca	881		
Db	17	ttctctctgttttctttagtattccatcagcagcttgcattgagcagcagctccca	76		
Oy	882	cttgaacacctgcgaataagcccgagagcgtgctcttcataatlttcaggtattgataa	941		
Db	77	ctttgaacacctgcgaataagcccgagagcgtgctcttcataatlttcaggtattgataa	136		
Oy	942	gaaaccgcgataagataacttttccacagcgtgagataataactgaatccatagaaaggac	1001		
Db	137	gaaaccgcgataagataacttttccacagcgtgagataataactgaatccatagaaaggac	196		
Oy	1002	tgtggaatttaaaatgttcttcttcataatccatcaagaacatccatccaaagattctaa	1061		
Db	197	tttggaatt	256		
Oy	1062	aggcttgaatcctcagaatcaatgaatctctgagagacagtcgcttgcgttctcaatgagcag	1121		
Db	257	aggcttgaatcctcagaatcaatgaatctctgagagacagtcgcttgcgttctcaatgagcag	316		
Oy	1122	tgggaagagtagcgtgagtcagcagcttctcagaaggttatatgatccgagatgagctttat	1181		
Db	317	tgggaagagtagcgtgagtcagcagcttctcagaaggttatatgatccgagatgagctttat	376		
Oy	1182	catgtgtgagtagaataagcatcagaagctt	1210		
Db	377	catgtgtgagtagaataagcatcagaagctt	405		
RESULT	6				
LOCUS	AK020318	872 bp	mRNA	HTC	05-JUL-2001
DEFINITION	Mus musculus adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230106F14, full insert sequence.				
ACCESSION	AK020318				
VERSION	AK020318.1	GI:12860872			
KEYWORDS	CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA, clone:libRIKEN full-length enriched mouse cDNA library clone:9230106F14.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 872)				
TITLE	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Methods in enzymology. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2 (bases 1 to 872)				
TITLE	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome research. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	11042159				
AUTHORS	3 (bases 1 to 872)				
TITLE	Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,				

TITLE	JOURNAL	MEDLINE PUBMED	REFERENCE AUTHORS
Samli, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kire, A., and Hayashizaki, Y.	Riken Integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer	Journal of Molecular Biology 205(3):913-917 (2000)	Reference authors
Functional annotation of a full-length mouse cDNA collection	Nature 409, 685-690 (2001)	5 (bases 1 to 872)	Adachi, J., Alzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carinoni, P., Fukuda, S., Fukushima, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Horii, F., Imotohata, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Otsuka, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Science Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsutsumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details.	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGGACATCTCCAGTAAATTAAATTAATTCGCCCGCCC 3'] cDNA was prepared by using triethanolamine thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGGACATCTCCAGTAAATTAAATTAATTCGCCCGCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid vector KS(+) after bulk excision from Lambda FLIC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.	Location/Qualifiers 1..872 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /db_xref="MGD:MGI:1911565" /db_xref="MGI:MGI:1924956" /clone="9230106F14" /sex="male" /tissue_type="epididymis" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 86..>872 /note="putative" /codon_start=1 /protein_id="BAB32070.1" /db_xref="GI:12860873" /translation="MANSEPRNGLOETNORGPLOEOVPKGNQAVPIEFRRADNINL DIVYLTITVIGCAALIRGVQVLSFWVTTRADOTTRIRKOFHSIIAODISWPDGSLS ICENLRNRDIDINKLGIDGKIPLMKNTSVTSFGSIFIGIVSLSTSKWSKLTVLVLSPIIL

BASE COUNT	262 a	178 c	187 g	245 t			
ORIGIN	MASSLCSRMIIISLTSKRLDYSRGNAAEALSSIQVTFNFGA*						
Query Match	19.5% Score 378.8; DB 12; Length 872;						
Best Local Similarity	76.0%; Pred. No. 1.4e-92;						
Matches	514; Conservative 0; Mismatches 117; Indels 45; Gaps 2						
QY	1 ttccgcttgcgcgtagagctgacatcaacatcatgatccctggtatctagctgacatcagc 60						
DB	197 TTCGCGTTTGGCAGTAATATGTGACATATTCCTCATATGACCTGGGATATTAGATCCAG 256						
QY	61 gtcaatgagcctgcgccttccctttaaagcactggtttttgagagaaatgagatcaactt 120						
DB	257 ATTAATATGGAGCCACCGCTTCCTTAATATGTCCTGGTTTATGAGAAATAGTATCATTTA 316						
QY	121 atagtagagctgctgtctgccaactaacaa----- 150						
DB	317 ATTTAATGGATGGCCTAGTATCAAACTTAACAACTTAATATACAACTGTTCTCAGACTCAA 376						
QY	151 -----tactcttcttcaggttgacccgttatatgttgagaaatagttgtgt 198						
DB	377 GAAAAGCTGATGAAAGATATCATTTGTTGTTGACTCTATATTATATTGGAATAGAGACAGCT 436						
QY	199 gctctgatttttggtttaacatacagaatttcccttggtagatataactgcaagacagacagc 258						
DB	437 GCCGTCATTTTGGCGATGTACAGATTTCTCTGGGTCATTAAGTCACGCCGGCAACCC 496						
QY	235 aagaggaatcgaaaaaagatttttcaatcagttttggcaagagacatcgctggtttgat 318						
DB	497 ACAAGATATCCGAAACAGATTTTTCATTCATTAATTTGGCAACAGACATCAGCTGGTTTGAT 556						
QY	319 agctctgacatcggtgaaacttaacactgcagatgac---agacatgacaaaatcagttgat 375						
DB	557 GCGAGTACATCTGTGAACTTAAACCCGCAATGACTGGTGACATCAACAACTCTGTGAT 616						
QY	376 ggtattgagataaagattgctctgttctgttttcaaaaacatgcttactcttcttgcattgacctg 435						
DB	617 GGTATTTGGAGATTAAGATAGTCCCTCTGATGTTTCAAGAACATATCTGGGGTTTCTATTTGGCCGTG 676						
QY	436 gcaagtgttggtygtaagagctgtaaaactcacccttagtgatctctatcaagctctctctt 495						
DB	677 GTGATTAAGTTTATATTAATAAAAGCTGGAACACTCTCCCTGGTGTCTGTCTACATCTCCTCTTC 736						
QY	496 ataaagagcttcagcgagcagcatgctctagatggtatcatctaatcttgacacgtagaata 555						
DB	737 ATTAATGGCTTCAATCGGCACTGTGCTCTAGAGATGATTAATCATTTGACCAAGCAAGAGCTG 796						
QY	556 agtgcctatccaaaagctgggctgtygagagaagaagcttctgcatcaatccgaacagtc 615						
DB	797 GATGCTTATCCAAAGCTGGGGCTGTGGCTGTAAGAAAGCTTGTCTATCAATCCAAACAGTC 856						
QY	616 atagaccttaaggccc 631						
DB	857 ACAGCCTTTGGAGCCC 872						
RESULT	7						
LOCUS	BF584668	894 bp	mRNA	EST			
DEFINITION	6020098406f1 NCI_CGAP_Co24 Mus musculus CDNA clone IMAGE:4218385 5',						
ACCESSION	BF584668						
VERSION	BF584668.1	GI:11658386					
KEYWORDS	EST.						
SOURCE	house mouse.						
ORGANISM	Mus musculus						
REFERENCE	Mammalia; Euteheria; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . Mammalian Gene Collection (MGC)						



```

Db      254  ACTTCCTGTCATCTCCAAAGAA-----TACTTATGACAAAGTGTCACTGTCTTC 306
Qy      835  ttatgtatcatcatagcagctatgtcatgtgagcagcagctccactcttgaacctc 894
Db      307  TTTTCCGTTTAAATGGACATTCAGTGTGGACAGGCTTCCAAATATTTGAAGCTTC 366
Qy      895  gaaataagcccgagagcctccttcatatcttcagagttatgaaagaccagata 954
Db      367  GCCATTCACGAGGACCTTATGACCTTCAAAATATTTGATATTAAGCCCAATATA 426
Qy      955  gataactcttcacagcttgatataaaccatgataacatagaagaaactgtgaattaa 1014
Db      427  GACAGCTTCAAAAGGTGGGCAACAACGACACATACATCAAGAAATCTGGAATTTAAG 486
Qy      1015  aatgttcttcaatcatcatcaagaacacatcatcaagaattctgaaagctgaaatc 1074
Db      487  AATATTCACCTGATACCTACCTCGAAAGAAAGTTCAGATCTTGAAGGCGCTCAATCTG 546
Qy      1075  agaattaagctcgagagagcagtcgcctgtgtcgtctcaatgagcagctggaagtag 1134
Db      547  AAGGTGAAGCGGAGACAGAGGCTGCTGTGGCAACAGTGGCTGTGAAAAGCCACA 606
Qy      1135  gtagcagcagctctcgagaggtatatagtaccgagatgagcttcatgtgtgagtag 1194
Db      607  ACTGTCCACCTGATGCAAAAGGCTCTACGACCCCTA-GATGGCATGTGATCGACGGA 665
Qy      1195  aatgacatcagagctttaaagtgcgcacatctcgagacacatatgagtggttagcaa 1254
Db      666  CAGGACATCAGAACCATCCAAATGTGAGTATCTGAGGAGATCAGGGGTGGAATCA 725
Qy      1255  -gagcctgtgtgttcgagagacacacatcagtaacaatacaagtatgagcagatgat 1313
Db      726  GGAACCTGCTGTGTTCGCCAACCCAGATCGCAGAAAATTCGTATGTGGGCGAGGAGTT 785
Qy      1314  gactgatgaagagatcg 1330
Db      786  GGTTCCTGTGTGATTTG 802

RESULT 9
Bg587938 780 bp mRNA EST 11-APR-2001
LOCUS     ESI489713 MHAM truncatula/Glomus versiforme mixed EST
DEFINITION library cDNA clone pMHAM-51L19 5' end, mRNA sequence.
ACCESSION Bg587938
VERSION    Bg587938.1 GI:13603002
KEYWORDS   EST.
SOURCE      Medicago truncatula/Glomus versiforme mixed EST library.
ORGANISM   Medicago truncatula/Glomus versiforme mixed EST library
REFERENCE   1 (bases 1 to 780)
AUTHORS     Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Uterback,T., Cho,J.
            and Fraser,C.M.
TITLE       ESTs from roots of Medicago truncatula after colonization with
            Glomus versiforme, 2001
JOURNAL     Unpublished (2001)
COMMENT     Contact: Harrison M.J.
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            Noble EST name: N387011e TIGR sequence name: MTDMM70TK More
            information is available at: http://www.medicago.org
            Seq primer: SKmod (CTA gaa gta gta gta CC).
FEATURES
Source      1..780
            Location/Qualifiers
            1..780
            /organism="Medicago truncatula/Glomus versiforme mixed EST
            library"
            /cultivar="Medicago truncatula genotype A17"
            /db_xref="taxon:119092"

```

```

/clone="pMHAM-51L19"
/clone_lib="MHAM"
/tissue_type="roots colonized with Glomus versiforme"
/dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
post-inoculation with Glomus versiforme. The library was
made from a mixture of RNA from each of these stages."
/lab_host="E. coli strain XL0LR"
/notes="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI. cDNA was prepared from polyA+ enriched RNA from
roots harvested at 10, 17, 22, 31 and 38 days
post-inoculation with Glomus versiforme. The cDNA was
directionally ligated into the UniZap XR vector from
stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-ass1st
helper phage and propagated in XL0LR cells."
BASE COUNT 234 a 156 c 184 g 206 t
ORIGIN

```

```

Query Match      12.0%; Score 233.6; DB 11; Length 780;
Best Local Similarity 58.7%; Pred. No. 7.2e-53;
Matches 404; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

```

```

Qy      863  ttgagcagcagcctccacttgaaccttcgcaatagcccgagagctgcttcata 922
Db      4  TGGGGCAATCTGCACCAGCATGGCTGCATTTACAAAGCTTACAGTGGCGCTCTAAGA 63
Qy      923  ttccaggtatctgaataaagccagatagataactttccaagcttgatataaac 982
Db      64  TTTCCGATTAATTGATCAACAGCTGTATAGTAAACAGTCAATCGATGTGGAAT 123
Qy      983  ctgaatccatagaagaactgtggaattaaagtcttcttcaattcatcatcaagc 1042
Db      124  TAGAGCACTTACTGACTGACTTGAACCTGAAAATGTGACCTCTTATCACTCAAGAC 183
Qy      1043  catctcaagattctgaaagctgtgaatcctcgaattgaatgctgagagacagctgc 1102
Db      184  CTGAAATTTCTGATCTCTCAATGATTTCTCTTGAGTGTTCCTGCGGAAGACCATATG 243
Qy      1103  tgtcgcgtctcaatgagcagtggaagagtagcagtagtccagctctgcagaggtatag 1162
Db      244  TAGTTGTGACAGCGGCTCTGGCAAGACACTGTGTTCTCTTATTTGAGAGATTTCTAG 303
Qy      1163  atccgagatgagcttcatcagtggtgagtagaagatgacatcagagcttcaatgtcg 1222
Db      304  ATCCAACTTCAGGACAAAGTATGTTGATGGGCGCATGACATTTAAACTTGAACCTTAAT 363
Qy      1223  attatcgagaccatattgagtggttagtcaagaagccttctgtcggagaccacata 1282
Db      364  GTTGAGGCAACAATAGAGCTAGTAGCCCAAGACCTGCTGTGTCACACGATTC 423
Qy      1283  gtaacatatcaagtagtagcagagatgagtagtcatgaaagagtagagagagcaaa 1342
Db      424  GAGAAATATTACTCTTGGAGAGCCTGATGCAACCAAGTGTGAGTTGAAGAAGTGTGA 483
Qy      1343  gggaagcaaatgcgtatgattatcatgagagtttccataaatttaacatgtgtg 1402
Db      484  GGGTTCTTATATGCTCATTTATTCATCATCAAAATTTCTTAAGGCTTTGAACCTCAG 543
Qy      1403  gggaagaaagagctcaaatgagtgagggcagaagaagagatcgcaatgtcgtcgct 1462
Db      544  GAGAAAGAGGACTTCAACTTTCTGGAGGACAAAACAAAGATATGCAATGCAAGGCA 603
Qy      1463  tagttcgaagcccaagatctgatttagatgagctacgtctgccttgatcagaa 1522
Db      604  TGTATAAATATCCAGCAATTTCTCTCCGATGATGAGCAACAAGATGATGACTGTGAAT 663
Qy      1523  gcaagtcagctgttcaagctgacgtga 1550
Db      664  CAGAAAAGCTGTGCAAGAAGCACTTGA 691

```







Db 428 TGGAGCTAGATGAGAGAAACCAACTGAAGAGATGATTTGATGCTGCTAAATGGC 369  
Oy 1350 aaatgcatatattatcatcaggagttcccaataattatataatgtaggagaa 1409  
Db 368 AAATGCACATATATTTATTTCTCATCTTCAGAACGATATGATTAAGTGAAGAAA 309  
Oy 1410 aggaagctcaatgtaggagagcagaacagagatcgcaattgctgctgcttagtc 1469  
Db 308 AGGAGCTCATTTATCAGAGAGACAAAACAAAGATTCTCTATGTCACGCTGCTTATAG 249  
Oy 1470 aaacccaaagattcgtatattagatgagcagtcgtccctggaattcagaagaagtc 1529  
Db 248 AAAACCTCAATTTTATTTACTTGATGATGATGATGATGATGATGATGATGATGAT 189  
Oy 1530 agctgtcaagctgacgagagagata 1558  
Db 188 GATGTGACAAAGCCTTGAAGAAAGCTA 160

RESULT 13  
LOCUS A2687805 880 bp DNA 14-DEC-2000  
DEFINITION ENT1J52TF Entamoeba histolytica Sheared DNA Entamoeba histolytica  
ACCESSION A2687805  
VERSION A2687805.1 GI:11824951  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica.  
ORGANISM Entamoeba histolytica.  
REFERENCE 1 (bases 1 to 880)  
AUTHORS Loftus B., Van Aken S. and Fraser C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
JOURNAL HMI:IMSS sheared DNA library  
COMMENT Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: entaeflgr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 30  
High quality sequence stop: 858.  
Location/Qualifiers  
1..880  
/organism="Entamoeba histolytica"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHOSt1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The V + I method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

BASE COUNT 247 a 167 c 122 g 344 t  
ORIGIN  
Query Match 11.3%; Score 218.6; DB 13; Length 880;

Best Local Similarity 58.7%; Pred. No. 9.4e-49;  
Matches 402; Conservative 0; Mismatches 274; Indels 9; Gaps 1;

Oy 883 ttgaacccctgcacatgcccagagagctgcttccatatttccaggtattgatatag 942  
Db 849 TTAATATATATATGCTACAGCACTTAACCTTCTGCATATATTAATTAACAACATATGATA 790  
Oy 943 aaacccagatagataacttccacagctgataataacatgataatcagaagaact 1002  
Db 789 ATCCAGATTTGATTTGATGATCTACACCTGGTGAATGTCACAGAGATGATGGAAT 730  
Oy 1003 gtgaaattaaatgcttcttccatataatcaatgaacacatcctaagaattctgaa 1062  
Db 729 ATTACATTGAGAAGATGTCATATTTGATATCCAAAGAACCAACTAACAATTCCTTGGT 670  
Oy 1063 gctcgaatcacaagataatgctgagagagacgctgcttgctgctcctaagcgt 1122  
Db 669 GCACTTGATCTTGAATAATTAAGAACCAACAGTTGATTAAGAGCATCAGCATGT 610  
Oy 1123 ggaagagtagtagtcacgctctcgcagaggtatataatccgagatgagcttalc 1182  
Db 609 GGTAAATCACTACTATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 550  
Oy 1183 atgctgagatgagagatcagacagagctttaaattgctgagcatatcagagacatagga 1242  
Db 549 AAATAGATGGAAGAAAGATTTAAGAGATTTAATTAATTAATTAATTAATTAATTAATTA 490  
Oy 1243 gtgtgtagcagagagcctgttctgctgagacacacatcagtaataatcagatagga 1302  
Db 489 TTAGTTGAGCAAGAACCATTTTGTTCATGATGATGATGATGATGATGATGATGATGAT 430  
Oy 1303 -----cgagatgagctgacatgagatgagatgagagagagagagagagagagagag 1353  
Db 429 GCTACAGATGAGAGAAACACCAACTGAAGAGATGATGATGATGATGATGATGATGATGAT 370  
Oy 1354 gcatgagatattatcatcagagagcttccataataattacatcagtaggagagagagag 1413  
Db 369 GCACATGAAATTTATTTTCATCTTCAGAGAGATGATGATGATGATGATGATGATGATGAT 310  
Oy 1414 gctcaatgagatgag 1473  
Db 309 GCTGATTTATCATCAGAGAGCAAAACAAAGAAATGCTATTTGACGCTGATTTGATTTGAAA 250  
Oy 1474 cccaagattcgtatttctgagagcagctgctgctgctgagagagagagagagagagag 1533  
Db 249 COTACATTTTATTTACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190  
Oy 1534 gtcaagctgacgtagagagagata 1558  
Db 189 GTACAAAGACACTTGAAAAAGCTA 165

RESULT 14  
LOCUS AM342644/c 575 bp mRNA 31-JAN-2000  
DEFINITION f181f06.x1 Sugano Kawakami zebrafish DRB Danio rerio cDNA clone  
2644067.3, similar to TR:Q91586 Q91586 MULTIDRUG RESISTANCE  
PROTEIN, mRNA sequence.  
ACCESSION AM342644  
VERSION AM342644.1 GI:6839010  
KEYWORDS EST.  
SOURCE zebrafish.  
ORGANISM Danio rerio.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Rasbora; Danio.  
REFERENCE 1 (bases 1 to 575)  
AUTHORS Sugano S., Kawakami K., Johnson S., Li F., Marra M., Eddy S.,  
Hillier L., Clifton S., Allen M., Gibson M., Jost S., Kucaba T.,  
Martin J., Pape D., Steptoe M., Underwood K., Theisgen B., Rutter  
E., Bowers T., Wylie T., Waterston R. and Wilson R.  
TITLE Mashu Zebrafish EST Project 1999  
JOURNAL Unpublished (1999)



```
OY 1054 attctgaagagctcgaatcagaattaagctcggagagacagtcgcttgcgctc 1113
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 ATTCTTGCTGGACTTGACCTTGAAATTAAGAAAGAGAAACAGTTCATTAGTAGGACA 201
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1114 aatggcagtcggagagtagcgttagtcgaagcttcgcagaggtatatagtaccgagat 1173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 TCAGGATGTGTAAGTCACTACATTAATTAAGTCCAAAGAGTCTATGATCCAGTTGGT 261
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1174 ggccttcatcgttggatggagaaatgacacagagcttaaatgtgcgcattacgagac 1233
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 GGAAAAGTAACATTAGATGGAAATGAAATTAAGAGAGTTGAACCTTGAAATGGTTAAGAAAT 321
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1234 caattggagtcgttaagcaagagccttcttgcggagaccacacagtaacaatc 1293
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CAATATAGGATTAGTTGACAAAGAACCTATTTTGTTGTCATGTACTATTAGAGAAATATT 381
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1294 aaglatggacgagatgag-----tgactgtagagagatggagagagcagcaag 1344
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 ATGCTTGAGCTAAGATGGAAACACCAACTGAGAGAGATGATGAATGTGCTAAA 441
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1345 gaagcaaatgcgtatgattatcatcgtgagttccataataattatcatctgttagg 1404
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 ATGCAAAATGCACATGAATTTATTCATCTTCAGAAAGCATATGATACAAATGTAGGA 501
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1405 gaanaagagagctcaatgagtcgagggcagaacagagagatcgcaatgctgcgctta 1464
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 GAAAGAGAGAGCTGCATTTATCGAGAGACAAACAAAGAAATGCTATTGCGACGTGCATTG 561
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1465 gttcgaaaccccaagatctcgtatttagatggagctacgctcgcctggtatcagaagc 1524
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 ATTAGAAACCTACATTTTATTTACTTGATGAAAGCTACATCAGCATTGATACACAAAGT 621
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1525 aagtcagctgtcagagctgcactgagagagata 1558
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 GAAAGATTGTACAAAGCAAGCACTTGAAAAAGCTA 655
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 22, 2002, 18:32:41  
Job time: 12254 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 18:46:40 ; Search time 293.81 Seconds  
(without alignments)  
1495.413 Million cell updates/sec

Title: US-09-873-409-16

Perfect score: 1940

Sequence: 1 ttccgcttgctgctgact.....taaatgctgacttatgga 1940

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/6C.COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/6D.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603.8	31.1	4646	1 US-08-181-471-2	Sequence 2, Appli
2	603.8	31.1	4669	2 US-08-752-447-1	Sequence 1, Appli
3	603.8	31.1	6505	2 US-08-793-610-5	Sequence 5, Appli
4	603.8	31.1	9318	2 US-08-793-610-6	Sequence 6, Appli
5	602.2	31.0	4669	6 5206352-3	Patent No. 5206352
6	597.4	30.8	4669	2 US-08-583-276-18	Sequence 18, Appli
7	594.4	30.6	4264	2 US-08-784-649A-1	Sequence 1, Appli
8	592.8	30.6	4264	2 US-08-784-649A-5	Sequence 5, Appli
9	568.4	29.3	4233	3 US-09-120-513-1	Sequence 1, Appli
10	568.4	29.3	4233	4 US-09-450-105-1	Sequence 1, Appli
11	258	13.3	2726	1 US-08-461-823-1	Sequence 1, Appli
12	243.2	12.5	4002	2 US-08-996-545-1	Sequence 1, Appli
13	243.2	12.5	4002	2 US-08-996-545-3	Sequence 3, Appli
14	243.2	12.5	4002	4 US-09-328-320-1	Sequence 3, Appli
15	243.2	12.5	4002	4 US-09-328-320-3	Sequence 3, Appli
16	230.8	11.9	4224	1 US-08-612-521-1	Sequence 1, Appli
17	219.4	11.3	4047	2 US-08-612-734B-1	Sequence 1, Appli
18	219.4	11.3	4800	2 US-08-612-734B-3	Sequence 3, Appli
19	189.6	9.8	3924	1 US-08-395-246C-1	Sequence 1, Appli
20	142.8	7.4	2061	4 US-09-061-764A-17	Sequence 17, Appli
21	141.2	7.3	1959	4 US-09-061-764A-4	Sequence 4, Appli
22	121.6	6.3	2376	1 US-08-394-880B-1	Sequence 1, Appli
23	116	6.0	3924	2 US-08-996-644-3	Sequence 3, Appli
24	116	6.0	3924	2 US-09-352-552-3	Sequence 3, Appli
25	116	6.0	3927	2 US-08-996-644-1	Sequence 1, Appli
26	116	6.0	3927	3 US-09-352-552-1	Sequence 1, Appli
27	105.4	5.4	5120	3 US-08-772-270A-6	Sequence 6, Appli

28	105.4	5.4	8370	2 US-08-488-706-1	Sequence 1, Appli
29	105.2	5.4	2244	4 US-09-061-764A-18	Sequence 18, Appli
30	104.4	5.4	4403765	4 US-09-103-840A-2	Sequence 2, Appli
31	104.4	5.4	4411529	4 US-09-103-840A-1	Sequence 1, Appli
32	103.6	5.3	6143	1 US-08-612-521-3	Sequence 3, Appli
33	100	5.2	7721	3 US-08-772-270A-14	Sequence 14, Appli
34	98.2	5.1	2407	2 US-08-895-522-2	Sequence 2, Appli
35	98.2	5.1	2407	3 US-08-195-391-2	Sequence 2, Appli
36	91	4.7	4403765	4 US-09-103-840A-2	Sequence 2, Appli
37	90.6	4.7	3909	1 US-08-232-537-1	Sequence 1, Appli
38	89.4	4.6	28804	2 US-08-592-874-1	Sequence 1, Appli
39	89.4	4.6	28804	3 US-09-096-942-2	Sequence 2, Appli
40	89.4	4.6	28804	3 US-09-096-867-2	Sequence 2, Appli
41	86.4	4.5	1201	3 US-08-961-083-189	Sequence 189, App
42	86	4.4	5889	1 US-08-463-092B-5	Sequence 5, Appli
43	86	4.4	5889	2 US-08-462-109A-5	Sequence 5, Appli
44	86	4.4	5889	2 US-08-460-907B-5	Sequence 5, Appli
45	86	4.4	5889	3 US-08-463-179A-5	Sequence 5, Appli

## ALIGNMENTS

RESULT 1  
US-08-181-471-2  
Sequence 2, Application US/08181471  
Patent No. 5641508  
GENERAL INFORMATION:  
APPLICANT: Li, Lingna  
APPLICANT: Lishko, Valeryi K.  
TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Thomas Fitting  
STREET: 12526 High Bluff Drive, Suite 300  
CITY: San Diego  
STATE: CA  
COUNTRY: USA  
ZIP: 92130  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/041,553  
FILING DATE: 02-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitting, Thomas  
REGISTRATION NUMBER: 34,163  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-792-3680  
TELEFAX: 619-792-8477  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 425..4267  
US-08-181-471-2









```
APPLICATION NUMBER: DE 195 03 952.1
: FILING DATE: 07-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PC7/EP95/03175
: FILING DATE: 10-AUG-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Berman, Richard J.
: REGISTRATION NUMBER: 39,105
: REFERENCE/DOCKET NUMBER: P1614-7007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)638-5000
: TELEFAX: (202)638-4810
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9318 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: DNA
: US-08-793-610-6

Query Match      31.1%; Score 603.8; DB 2; Length 9318;
Best Local Similarity 65.6%; Pred. No. 1.4e-173;
Matches 916; Conservative 0; Mismatches 472; Indels 9; Gaps 2;

QY 162 caggttcacccctgtatattgttggaatagtgctgctccttgattttggtatataca 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2111 CAGGATATCCCTATATATACATGGAATGCTGGGCTGCTGCTGCTGCTGCTACATCA 2170

QY 222 gatttccttgaggattaaatcgacagacagacaaaggattcgaaaaaatctttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2171 GGTTCATTTTGGGCGCTGCGAGCTGGAAGACAAATACAAAATTTAGAAACAGTTT 2230

QY 282 tcatcagttttgacacagacacatcgctggtttgatagtctgacacatcggtgaactaa 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2221 TCATGCTATATATGCGACAGGAGATAGCGTGGTTGATGACGATGTTGGGGAGCTTAA 2290

QY 342 cactcgcatcagaacattgac--aaatcagtgatgattgagataaagattgctct 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2291 CACCGGACTTACAGATGATGCTCTTAAGATTAAATGAGTTATGGTGACAAATTTGGAAT 2350

QY 399 gtttttcaaaaacatgctatcttcttgattggtcgctgcaagtgtgttgaaaggctg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2351 GTTCTTTAGTCAATGGGACAAATTTTTCACCTGGGTTTATAGTAGGATTACACGTGTTG 2410

QY 459 gaaactcaacctagtgactctatccaagctctcctcttaaatgagctcaagcgacagctg 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2411 GAAGCTAACCTTGTGATTTTGGCCATCAGTCCCTGTCTTGAGACTGTGACGTGCTGTCTG 2470

QY 519 ttctagatgltcatctatctgaccagtaagaalttaagtgctctatccaaagctgggac 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2471 GGCAGAAAGTACTATCTTACTTACTGATTAAGAACTCTTAGCGGTATGCAAAAGCTGAGC 2530

QY 579 tgtggagaaagaagctgtgatactcgaacagctcatagctctttagggcccaagagaa 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2531 AGTAGCTCAAGAGCTGTGGAGCAATTAAGACTGTGATTTGATTGGAGGCAAAAGAA 2590

QY 639 agaactcaaaaggtatcacagaatctcaaaagatgcaagaatttctggcataaaagac 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2591 AGAAGCTTAAAGGATACACAAAATTTAGAGAAAGCTTAAAGAAATTTGGGATTAAGAAC 2650

QY 699 tatagctcaaaagtgctcttggtgtgtgtaactcttataatgaaatgaaactatgagact 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2651 TATATACACCAATATTTTATAGGTGCTGCTTCTGCTGATCTATGCAATCTTATGCTCT 2710

QY 759 tgccttttgatagaaactcttgattcttaaatgagaactgagactgatactatggagac 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2711 GCCCTTCTGGATGGGACACCTTGCTCTCAGGGA-----ATATTTATTTAGACA 2764

QY 819 tgtcttgctgttctcttgatgtaatcatagacagttatgcatctgagcagcagctcc 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2765 AGTACTCACTGATATCTTTCTGATTTAATGGGCTTTTAACTGTGGACAGGCACTCTCC 2824
```

```
QY 879 tcaacttgaaccttcgcacatagcccgaggagctgccttcaatatttccaagttatga 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2825 AAGCATTTGAAGCATTTTGGAAATGCAAGAGGACGACTTATGAAATCTTCAATATATGA 2884

QY 939 taagaatacccagtatagataaactttccacagctgatatataaactgaaatccatagaag 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2885 TATATAGCCCAATATTGACAGCTATTTCGAAGAGTGGGCAACCAAGATTAATATTAGGG 2944

QY 999 aactgtgaaatttaaaaatglttcttcaattatccatcaaaagaccatctacaagattct 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 2945 AATATTTGGAATTCAGAAATGTTCACTTACGTACCTAACCATCTCGAAAAAGAGTTAAATCT 3004

QY 1059 gaaaggtctgaatctcagaataaagtctgagagagacagtcgccttgctgcgtcgaatg 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3005 GAAGGCGCTGAACCTGAAGGTCCAGAGTGGGAGACGGTGGCCCTGCTTGGAACAGTGG 3064

QY 1119 cagttgggaagtaagtgatgcagctctctgcaagggtatataatgacgagatgctt 1178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3065 CTGTGGAGAGACACACAGTCCAGCTGATGCAAGGCTCTATGACCCACAGAGGGGAT 3124

QY 1179 tatcatgtgtgataagaaatgacatcagagctttaaatgtgcgcatltaacgacacatat 1238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3125 GGTCAAGTGTGATGAGACAGATATTAGACACATAATGTAAGGTTTCTACGGGAAATCAT 3184

QY 1239 tggagtggttagtcaagagcctglttgttcggggaccacatcaagtaacaatatcaagta 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3185 TCGTGTGGAGTACAGGAACCTGATTTGTTTGCACACAGATAGCTGAAACCATTCGCTA 3244

QY 1299 tggagcagaatgatagtacatgaaagagaatgagagaagacagaaggaaagcacaatgcgta 1358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3245 TGGCGCTGAAATGTCACCATGATGATGATGAGAAAGCTGTCAGAGAACCAATCCCTA 3304

QY 1359 tgatttalcataagagtttccctaataatcttaataatctgtaggggaaagagagctca 1418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3305 TCAGCTTTATCAGTAAGAACTGCTCATTAATTTTGCACACCTGTTGGAGAGAGAGGGCCCA 3364

QY 1419 aatgagtgtagggcagaacaagaagatcgcaattgctgctgcttctgctgaaccccaa 1478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3365 GTTGAAGTGTGGGCAAGAACAGAGATCGCATTCGACGTGCCCTGTGCGAAACCCCAA 3424

QY 1479 gatttcgatttagatgaagctacgtctgcccgtgattcagaagaagaactgacgtgtca 1538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 3425 GATCTCTCTGCTGATGAGGCCACGCTGACGCTTTGGACACAGAAAGGAAAGCACTGTTCA 3484

QY 1539 agctgcacttgagaag 1555
    ||||| ||||| |||||
Db 3485 GGTGGCTGTGATTAAGG 3501

RESULT 5
5206352-3
: Patent No. 5206352
: Applicant: Ronlison, Igor B.; Pastan, Ira H.; Gottesman,
: Michael M.
: TITLE OF INVENTION: COMPOSITIONS FOR CLONES RESISTING DNA
: SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
: NUMBER OF SEQUENCES: 4
: CURRENT APPLICATION NUMBER:
: APPLICATION NUMBER: US/07/622,836
: FILING DATE: 24-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 892,575
: FILING DATE: 01-AUG-1986
: APPLICATION NUMBER: 845,610
: FILING DATE: 28-MAR-1986
: SEQ ID NO:3
: LENGTH: 4669
5206352-3

Query Match      31.0%; Score 602.2; DB 6; Length 4669;
Best Local Similarity 65.5%; Pred. No. 3e-173;
```

Matches 915: Conservative 0; Mismatches 473; Indels 9; Gaps 2;				
OY	162	caagtgaccctgattatgttgaaatggtgtgctgaccttattgttgcatacaca	221	
Db	760	caagtgaccctgattatgttgaaatggtgtgctgaccttattgttgcatacaca	819	
OY	222	gatttcctgtgtgattataaactgcaacacgacgacgaagaggtatcgaaacagtttt	281	
Db	820	ggttcatatttggctgtgctgacgtgaagacacaataacacaattgaacacagtttt	879	
OY	282	tcaattcaagtttggcacaagacatcgctgtgttgatgctgtgacatcggtgaactaa	341	
Db	880	tcatgtcataatgcaacgaagagatagcgtgtgttgatgctgtgacatcggtgaactaa	939	
OY	342	cactgcgatacagacattgac---aaatcagatgattgattggaataagattgtctc	398	
Db	940	caaccgactacacgatagttctcctaagattacaaattattggtgacaaatgtgaat	999	
OY	399	gttgattcaaaaacatgtctacttcttcgattggcctgacgtgtgtgtgaaaggctg	458	
Db	1000	gttccttcagtcacatggcaacatttccactggtgttatagtagatttaccagtggtg	1059	
OY	459	gaactacacctagtgactctatccacgtctcctctataatagcttcacgacgacatg	518	
Db	1060	gaagctaacctctgtgattcttgccacacgtctcctcttggaactgtcagctgtgtctg	1119	
OY	519	tctcagatgtcatctcattgaccagtaaggaattagtcctcattccaaagctggggc	578	
Db	1120	ggcaaaagatcatctcatttactgataaagaactcttgacgtatgcaaaaagctggag	1179	
OY	579	tgtygcagaagaagctctgtcatcatccgacacgacatagcctttagggccagagaa	638	
Db	1180	agtagcgaagaagctctgtcgaacatgaagacgtgtgatttgagagacaaagaa	1239	
OY	639	agaactcaaaagatatacacgaatccaaagatgcaaaagatttggcataaaaggac	698	
Db	1240	agaacttgaaggtacacaacaaatctagaagaagctaaagaattgggataaagaagac	1299	
OY	699	tatagctcaaaagtgctctctgtgtgctgtgacttcttataatgaaactatgac	758	
Db	1300	tattacagccaatattcttaagtgctgcttctgtgtgtctatgacttattgctct	1359	
OY	759	tgctttctgtatggaacctcctctgattcttaatggagaacctgtgataaccatcggaac	818	
Db	1360	ggcctctgtatggagacacctgtgtcctctcagggga-----atacttattggaca	1413	
OY	819	tgctctgtgttcttctttagtgaatccatagacgatttgatcttgagacagacgtccc	878	
Db	1414	agtaactgacttattcttctgtatataattggggctttagtgttgacagacgtctcc	1473	
OY	879	tcaattgaaacctgcgaatagccgagagacgtgaccttataatttccaggtatgta	938	
Db	1474	aagcatggaagcatttgaaatgcaagagggagcgcttatgaaatcttcaagatattga	1533	
OY	939	taagaacaccgatatagataacttccacagctgatatataaactgtatccatagaag	998	
Db	1534	taataagccaagatttgacagctatcgaagagtgycacaacacgaataatlaaggg	1593	
OY	999	aacgtgtgaatttaaaatgttcttcaattatccatcaagacatctatcaaatctc	1058	
Db	1594	aaatttgaattcgaatgttccattacatccatccgaaaaagaagtlaagctctc	1653	
OY	1059	gaaggtctgatactcgaatlaagctcggagagacagctgcctgtgctgcgtcacaag	1118	
Db	1654	gaagggcctgaaactggaaggtgacagagtgggcagacagctggttggaacagtg	1713	
OY	1119	caatgggaagagtaagtgagtcacagctctcgaaggttatatgacgagatgagctt	1178	
Db	1714	ctgtgggaagagacacaacagctcagctgtagagaggtctatgacccacagagggag	1773	
OY	1179	tatcatgtgtgataagatgaatcagagctttaaattgtggcgttatctcgagacatat	1238	
Db	1774	gttcagtgctgattgacagagatattagacataatglaaggttcttaccaggaatacat	1833	

OY	1239	tggagtgttagtcaagagcctgttctgttcggaccacacatcaagtaacatacaagta	1298	
Db	1834	tggtgtgtgagtcaggaacactgtattgtttgcaccacagatagctgaaacattcgcta	1893	
OY	1299	tggacagagatgattgtcatgataagaagtgagagagacgacgaaggaacaaatcgta	1358	
Db	1894	tggtcgctgaaatggtccacatgtgataagattgaaagcgttcacgaagacaaatgcta	1953	
OY	1359	tgaattatcatgagatttctcctaataattatcatattgtgtgaggggaaaaaggactca	1418	
Db	1954	tgaattatcatgaaactgtcctcataatttgacacacctgtgtgagagagagggccca	2013	
OY	1419	aatgagtggagggcagaacacagagatcgcaattgtctgtgcttatttgaaaccccaa	1478	
Db	2014	gttgagtgtgtggcagaagacagagatgagatcgcattgtcagctgtgttctgcacccaa	2073	
OY	1479	gattctgattttagatagagctacgtctgcctctgattcagaagaacagtaagctgtca	1538	
Db	2074	gattcctctgtgtatgtagggcagcagctcagccttggaacagaaagcgaagctgttca	2133	
OY	1539	agctgcactggagaag	1555	
Db	2134	gtgtgctctgataag	2150	

RESULT 6  
 US-08-583-276-18  
 ; Sequence 18, Application US/08583276  
 ; Patient No. 5837536  
 ; GENERAL INFORMATION:  
 ; APPLICANT: McDonagh, Kevin T.  
 ; APPLICANT: Nienhuis, Arthur  
 ; APPLICANT: Tolstoshev, Paul  
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
 ; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Carella, Byrne, Bain, Giffillan,  
 ; ADDRESSEE: Cecchi & Stewart  
 ; STREET: 6 Becker Farm Road  
 ; CITY: Roseland  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07068  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch diskette  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: DW4.V2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/583,276  
 ; FILING DATE: 05-JAN-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/332,444  
 ; FILING DATE: 31-OCT-1994  
 ; APPLICATION NUMBER: 07/887,712  
 ; FILING DATE: 22-MAY-1992  
 ; INFORMATION FOR SEQ ID NO: 18:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4669 bases  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: singular  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Genomic DNA  
 ; DESCRIPTION: Genomic DNA  
 ; US-08-583-276-18

Query Match 30.8%; Score 597.4; DB 2; Length 4669;  
 Best Local Similarity 65.3%; Pred. No. 8.6e-172;

Matches	912:	Conservative	0:	Mismatches	476:	Indels	9:	Gaps	2:
OY	162	caggttgaacctatatagttgaataagtggtgcttgatcttgggtacata							221
Db	760	CAGGTAGCCCTATTATATACATGGAAATGGTGGGCTGCTGCTGCTTACATCA							819
OY	222	gattcccttggtgattataactgacgcacgacccaaggaggttcgaagaagtttt							281
Db	820	GGTTTCATTTTGGTCCCTGGCAGCTGGAAACAAATACAAAATTAACAATTTT							879
OY	282	tcattcagttttgacacagacatcgctggtttgatagcgtgacacgctgtaactaa							341
Db	880	TCATGCTTATATGGACACAGAGATAGCTGGTTGATGTGACAGATGTTGGCAGCTTAA							939
OY	342	cactgcacatcagacacatgac---aaaatcagatggttatttggagataagattgctct							398
Db	940	CACCCGACTTTCACATGATGATCTCTAATGATTAATGAACTTATTTGGTGCAAAATTTGGAAT							999
OY	399	gttgttcaaaacatgctatcttcttgatgctgctgacgttggttggtggaaggtctg							458
Db	1000	GTTCTTTGATCAATGGACAAATTTTTCAGTGGTTTATAGATTAATTAACAGTGGTTG							1059
OY	459	gaactcaacctagtgactatccacgctctcctcttaataagcttcagcgcagacatg							518
Db	1060	GAAGCTAACCTTGTGATTTTGGCCATCAGTCTGTCTTGAGACTGTCAGCTGCTG							1119
OY	519	ttctagagatgcatcattcattgacccaagaaatlaagtgctattccaaagctgggac							578
Db	1120	GGCAAAAGTACTATCTTATCTTACTGATGAAGAACTTACGCTTACGATGCAAAAGCTGAGAC							1179
OY	579	gtgtgcagaagaagttctgtcatcaatccgcgaagctcatagcctttagggcccaagaa							638
Db	1180	AGTAGCTGAAGAGGCTTGGCAGCAGATTAAGAACTGTGATTTGATTTGGAGCAAAAGAA							1239
OY	639	agaactcaagaagatatacagaagatctcaagaatgcaagagatttggcataaaaggac							698
Db	1240	AGAACTTAAAGGTRACACAAAATTTAGAAAGAGCTTAAAGAAATTTGGATTAAGAAAGC							1299
OY	699	tatagctcaaaagtgctctgtgtgctgtgacttcttcttaagaaatggaactatgact							758
Db	1300	TATTAACACCAATATTTCTATAGGTGCTGCTTCTGCTGATCATGCAATCTTATAGCTCT							1359
OY	759	tgcttttggtatggaacctcttgatcttcaatgagaacctgataatacatctgggac							818
Db	1360	GGCCTTCTGGTATGGGACACCTGGTCTCTCAAGGGA-----ATATCTTATTTGGACA							1413
OY	819	tgctcttgcttcttctcttagtgtaatccatagcagttatgcatgagagcagagctcc							878
Db	1414	AGTACTCACTGATTTCTTTCTGTATTAATGGGCTTTTGTAGTGTGGACAGGCAATCTCC							1473
OY	879	tcactttgaaacctcgcaatagcccgagagctgcttccatatttccaggttattga							938
Db	1474	AAGCATTTGAACATTTGCAAAATGCAAGAGAGACAGCTTATGAATCTTCAAGATTAATGA							1533
OY	939	taagaagaacctatagataactttccacgctgataataaactgatacatcagaagg							998
Db	1534	TAAATTAAGCAAGTATTTGACAGTATTTGAAAGTGGGCAAAACAGATTAATTAAGGG							1593
OY	999	aacttggaatttaaaatgttcttccaatcatcacaagacatcatcaaatct							1058
Db	1594	AAATTTGGAATTTGCAAAATGTTCACTTCAAGTTACCCATCTGCAAAAAGAAATTAAGATCTT							1653
OY	1059	gaaagctgaactcagaatlaagctcgtgagagacagctgcgcttgctgcctcaatg							1118
Db	1654	GAAGGAGCTGAAGCTGAAGGTGACAGAGTGGGACAGGACCCCTGGTGTGAAGAACAGTGG							1713
OY	1119	cagtgaggaaagtaagcgtgagctcctcgtgagaggttatatgattccggaatgagctt							1178
Db	1714	CTGTGGGAGAGACAAACAGTCCACCTGATGACAGGCTCTATTAAGACCCACACAGAGGGGAT							1773
OY	1179	tatcatgttgatgaaatgatacatcagagctttaaattgctgcgcatattatcgagacatat							1238
Db	1774	GGTCAAGTGTTCATGACAGGATTAAGACCAATTAATGTAAGGTTTCTTAAGGGGAATCAT							1833

OY	1239	tggagtggttagtcaagagccgtgttctgtcgggaccacatcagtaatacatataca							1298
Db	1834	TGGTGTGGTGAAGTCAAGAACTGTATTTTGGCCACACAGATGCTGAAAATTCGCTTA							1893
OY	1299	tggacgagatgtagtgcagtgaagaagaatggagagacgaagaaggaaagcaatgcgta							1358
Db	1894	TGGCCGTGAATATGTCACCATGTGATGAGATTGAGAAAGCTGTCAAGAGAACCAATGCCA							1953
OY	1359	tgatttcaatgagatcttccataaatttaatacatatgtaagggaagaaagagctca							1418
Db	1954	TGACTTTATCAATGAAGACCTCTCATTAATTTGACACCCCTGTTGGAGAGAGAGGCCCA							2013
OY	1419	aatgagtgagggcagaacaaagagatgcgaattgctgtgcttaagttcgaaaccccaa							1478
Db	2014	GTTGAGTGGTGGGACAGAACAGAGATGCGCATTTGACAGTGCCTGTGTGCAACCCAA							2073
OY	1479	gattctgatttagatgaagctacgtctgccttgatcattcaagaagaagcaagctgtca							1538
Db	2074	GATCTCTCCTGCTGATGAGGCCACGCTGACACCTTGGACACAGAAAGGAGCATGTCA							2133
OY	1539	agctgcactgagaagg							1555
Db	2134	GGTGCCTCTGATTAAG							2150

RESULT 7  
 US-08-784-649A-1  
 ; Sequence 1, Application US/08784649A  
 ; Patent No. 5830697  
 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sikic, Branimir I  
 ; APPLICANT: Chen, Gang  
 ; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
 ; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; STREET: 2200 Sand Hill Road  
 ; CITY: Menlo Park  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94025  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/784,649A  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Sherwood, Pamela J  
 ; REGISTRATION NUMBER: Reg. No. 5830697 36,677  
 ; REFERENCE/DOCKET NUMBER: 06037/007001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-322-5070  
 ; TELEFAX: 415-854-0875  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4264 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-08-784-649A-1

Query Match 30.6%; Score 594.4; DB 2; Length 4264;  
 Best Local Similarity 64.7%; Pred. No. 6,7e-171;  
 Matches 902; Conservative 0; Mismatches 486; Indels 6; Gaps 1;

```

Oy 162 caggttgaccctgataatgtaagtgctgcttgatatttggtacataca 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 477 CAGGTATGCTTATTTATAGTGAATGTGCTGGGCTGCTGCTGCTTCAATTA 536
Oy 222 gatttccttgatataactcagacagacagcaagaagattcgaatacaagtttt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 537 GGTTCATTTTGGTGGCTGGCAGCTGGAACAATAACAAAAATTAAGAAAGTTT 596
Oy 282 tcatcaagtttgacagacacgcgcttgatagctgacacgcgtgaactaa 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 597 TCAATGCTATATATCGACAGAGATAGGCTGTGATGATGACAGATTTGGAGCTTA 656
Oy 342 cactgcacatgacagacatgacaaatcagtgatgataagaattgctctgt 401
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 657 CACCCGACTTACAGATGATGCTCAAGATTAAATGAAGGATTTGATGACAAATTTGGAAT 716
Oy 402 gttcaaaaatgctacttttcgattgcccgcgagctggttggtagaaggtcgaa 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 717 GTTTCAGTCATGACCAATTTTCACTGGGTTATAGTAGGATTTACACGTGGTTGAA 776
Oy 462 actcacccatgactcactacacgctcctctataatgcttcacgcgacatgttc 521
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 777 GCTAACCTTGTATTTTGGCCATCAGTCTGTTTGACATGCTGCTGCTGCTGGC 836
Oy 522 taagatggtcatcattgacagtaagaatgaatgctatccaaagctgggagct 581
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 837 AAGATACATCTTCTTATTTACTGATAAGAACTTTAGCTGATCAAAAGCTGAGCAGT 896
Oy 582 ggcagagaagaagcttctgcatcaatccgaacagcatagcctttagggccagagaga 641
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 897 AGCTGAAGAGGCTTTGGCAGCAATTAAGACTGATTCATTTGAGACAAAAAGAAAGA 956
Oy 642 actcaaaatatacaagaatcacaagatgcaagatttgcaataaaagactat 701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 957 ACTTGAAGGTACAAACAAATTTAGAAAGAACCTTAAGATTGGATTAAGAAAGCTAT 1016
Oy 702 agcttcaaaagtgctctggtgctgctgcttcttcaatgaatgaactatgactgc 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1017 TACAGCCAAATTTCTATAGTGTGCTGCTTCTGCTGATGATGATGATGCTGCTGGC 1076
Oy 762 ttttgataggaaccccttgatcttcttaatggaacccgatatccatcgagactgt 821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1077 CTTCTGTATGGGACCACTTGTGCTCTCAGGGGA-----AATTTCTATTTGCAAGT 1130
Oy 822 tcttgctgttcttgatgataccatagacatgatttgacagacagcctcctca 881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1131 ACACACGATTTCTTCTGATTAATTGGGCTTTAGTCTTGACAGGATCTCCAG 1190
Oy 882 cttgaaacctcgcaatagccgagagctgcttccatatttccaggttattgataa 941
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1191 CATTGAAGCATTTGCAATGCAAGAGAGAGAGCTTATGAATCTTCAAGATTAATTGATTA 1250
Oy 942 gaaacccagatagataacttttccacagctggtatataaactgaatccatagaagaac 1001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1251 TAAGCCAGATTTACAGCTATTTGGAAGAGTGGGACAAACCAATATATTAAGGGGAA 1310
Oy 1002 tctgaaatttaaaaatgttcttcaattatccatcaagacacataccaagattcgaa 1061
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1311 TTTCGATTTGCAATGCTTCACTTACCTTACCTTCCGAAAGAGTAAAGATCTTGA 1370
Oy 1062 aggtctgaatcagaaatgaagctgagagacagctgcttgctgctcaatgagcag 1121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1371 GGGCTGAACCTGAAGGTGAGAGTGGGACAGCGTGCCCTGGTTGGAACAGTGGCTG 1430
Oy 1122 tgggaagaagtaagtaagcagctctgcagaggttatataatccgagatgtgcttat 1181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1431 TGGGAAGAGCAACACAGTCCAGCTGATGACAGAGCTATATACCCACAGAGGGGAGTGT 1490
Oy 1182 catgtgtagatagatagatcagagctttaaattgctgagattatcgagacatattgg 1241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1491 CAGGTGATGAGACAGGATTTAGGACCATTAATGAAGTTTCTACGGGAAATCAATTGG 1550
Oy 1242 agtggtagtaagaagctgtttcttcgagacacacatcagtaacaataatcaatgag 1301

```

```

Db 1551 TGTGTAGTCAGAGAACCTTATTTGTCACCCAGATAGCTGAATAACATTCGCTATGG 1610
Oy 1302 acgagatgatgactgatgaagagatgagagacagcaaggaagcaaatgctatga 1361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1611 CCGTGAATAATGTCCACTTGATGATGATTAAGAAAGCTGTCAAGAACCAATGCTTGA 1670
Oy 1362 ttatcatgaggttccctaataatlaacatggtgaggggaaaaagagctcaat 1421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1671 CTTATATCAAACTGCTCATATAATTTGACACCCCTGTTGAGAGAGAGGGGCCAGTT 1730
Oy 1422 gagtggagggcagaagaagagatcgcaattgctcgtgcttggttcgaaccccaagat 1481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1731 GAGTGTGGGCAAGACAGAGATGCGCATTTGACGTGCCCTGTTGCAACCCCAAGAT 1790
Oy 1482 tctgatttagatgagctcagctgcccctgagattcagaagaacagctgttcaagc 1541
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1791 CTTCTCTGTGATAGGCCACAGCTGAGCTTGACACAGAAAGCAACAGATGCTTACGT 1850
Oy 1542 tgcactggagaag 1555
    ||||| ||||| |||||
Db 1851 GGTCTGTGATAGG 1864

```

```

RESULT 8
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: SIKIC, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-784-649A-5

```

```

Query Match 30.6%; Score 592.8; DB 2; Length 4264;
Best Local Similarity 64.6%; Pred. No. 2,1e-170;
Matches 901; Conservative 0; Mismatches 487; Indels 6; Gaps 1;
Oy 162 caggttgaccctgataatgtaagtgctgcttgatatttggtacataca 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 477 CAGGATGCTATATATACAGTAGGATTTGGTCTGGGCTGTGCTGTCTTACATCA 536
OY 222 gatttccttgygaltataactgcgcagcagacagcaagagattcgaaacagtttt 281
Db 537 GGTTCATTTTGGTCCCTGGCAGCTGGAGACAAATATACAAATTTAGAAACAGTTTTT 596
OY 282 tcaatcagtttggcagagacatgcgtgtgtatagctgtgacatgctgtaactaa 341
Db 597 TCATCTTATATGCGACAGAGATAGCGCTTGTGATGCGACGATGTTGGGAGCTTAA 656
OY 342 cactgcaltgacagacatlgacaaaacagtgatgtatgttgagatbaagattgtctgt 401
Db 657 CACCGGACTTACAGATGATGCTCCAGATATATGAAGAAATTTGGACAAATTTGAAAT 716
OY 402 gtccaacacatgctacttttctgattgctgcgcagctgtgtgttgtaagagctgaa 461
Db 717 GTTCCAGTCAATGAGCAATTTTTCACGTGGTTTATAGATGATTTACAGTGTTGGAA 776
OY 462 actcacccatgactctatccagctctcctcttataagttcagggcagcagatgttc 521
Db 777 GCTAACCCCTTGATGATTTTGGCCATCATGCTCTTCTTGAGACTGTCAGCTGTCTGGC 836
OY 522 taggatgtlcatlcatlcatlgaaccagaaatlaagtgcctatccaaaagctgagctgt 581
Db 837 AAAGATACTATCTTCTTACTGATGAAGAACTTAAAGCTATGCAAAAGCTGGAGCACT 896
OY 582 ggcagagaagtcctgtcatcaatccgaacagtcataagcctttagggccagagaaga 641
Db 897 AGCTGAAAGAGCTTGGCAGCAATTAAGAACTGTGATTTGCAATTTGGAGCAAAAGAA 956
OY 642 actcaaaagatatacagaagaatcgaagaatgaagaatlttgcatataaagaactat 701
Db 957 ACTGTAAGAGTACAAACAAATTTTGAGAAAGCTTAAAGAAATTTGGATTAAGAACTAT 1016
OY 702 agcttcaaaagtcctctgtgtgtgtgtactcttataagttgaatggaacctatgactgtc 761
Db 1017 TACAGCAATATTTCTATAGTGTGCTGCTTCTGCGATCTATGCACTTATGCTGTGCG 1076
OY 762 ttttctgtatggaacctcctgtatcttcttaatgagaacctgatatatacactgagactgt 821
Db 1077 CTTCGTGGTATGGGACCACTTGTGCTCTCAGAGGGA-----ATAATTTATGGACAACT 1130
OY 822 tctgtctgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 881
Db 1131 ACTGCTGATTTCTTTCTGCTTATTAATTTGGGCTTTTACTGTTGGACAGGCACTTCCAAG 1190
OY 882 ctttgaaccttcgcgaatagcccgaagagctgtccttctatcttccagagttatgataa 941
Db 1191 CATTTAGAGATTTGCAAAATGCAAGAGGACGCTTATGAATCTTCAACATATATTGATTA 1250
OY 942 gaaacccagtatagataacttcttccacagctgatatataaactgaaatcatalgaagaac 1001
Db 1251 TAAAGCAAGTATGACAGCTATTTGCAAGAGTGGGCAACAAACAGATAATTAAGGAGAA 1310
OY 1002 tgttgaatttaaaatgttcttcttcttcttcttcttcttcttcttcttcttctt 1061
Db 1311 TTTGGAATTCAGAAATGTTCACTCACTGATTAACCATCTGCAAAAAGAAATTAAGTTCTTGA 1370
OY 1062 aggtctgaatcgcgaatlaagtctgtgagaagagctgcgctgtgtcgtcctcaatgacg 1121
Db 1371 GGGCTGAAACCTGGAAGTGCAGAGTGGGCAAGCGTGGCCGTGTGGAACAGTGGCGTG 1430
OY 1122 tgggaagagtagcgttagtctcagctcttcagagagttatataatgacgagatgacttat 1181
Db 1431 TGGGAAGACACAAACAGTCCACCTGATGCAAGAGGCTCTATGACCCACAGAGGGGATGTGT 1490
OY 1182 catgtgtgagtagaatgacatagagcttaaatgtgagagcttatcagagacatattgt 1241
Db 1491 CAGTGTGATGACAGGATATTAAGGACATTAATGTAAGGTTTCTAAGGGAATCAATTGG 1550
OY 1242 agtgglttagtcaagagctgttctgttcgggaaccacatcagtaatacaatacaagtatgt 1301
Db 1551 TGTGGTGAAGTCAAGAACTGTATTTGTTTCCACACAGATAGCTGAAGAAACATTTGCGTATGG 1610

```

```

OY 1302 accagatgatgtgactgatgaagagatgagagcagcagcaaggaagcaaatgctatga 1361
Db 1611 CCGTGAATATGTCACCATGATGAGATTGAGAAAGCTGTCAAGAAAGCCAAATCCCTATGA 1670
OY 1362 ttatcatgtaggtttccataataatlaatacatgtgtgtagggggaagagagctcaat 1421
Db 1671 CTTTATCATGAAGAACTCCCTCATTAATTTTGACACCTGTGGTGGAGAGAGAGGGGCCAGTT 1730
OY 1422 gagtggagggcagaacagagagatgcgaattgtctgtcctttagtcgaaccccaagat 1481
Db 1731 GAGTGTGGGCGAAGCAAGATGCGCATTTGCACTGCTGCTGTTCGCAACCCCAAGAT 1790
OY 1482 tctgatttagttagagctacgtctgccttggatttcagaagaagctcagctgttcaagc 1541
Db 1791 CCTCCTGTGATGAGGCCAGTCAAGCTTTGGACACAGAAAGCAAGTGTTCAGGT 1850
OY 1542 tgcactgagaag 1555
Db 1851 GCGCTGATTAAG 1864

```

## RESULT 9

US-09-120-513-1  
Sequence 1, Application US/09120513

Patent No. 6025160

GENERAL INFORMATION:

APPLICANT: Brun, Kimberly

APPLICANT: Chenerly, Richard

APPLICANT: Ellens, Harna

APPLICANT: Field, John

APPLICANT: Yue, Lin

TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE

TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR12 AND

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY:

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/120,513

FILING DATE: 22-JUL-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: King, William T

REGISTRATION NUMBER: 30,954

REFERENCE/DOCKET NUMBER: GP50008

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5015

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4233 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

US-09-120-513-1

Query Match 29.3%; Score 568.4; DB 3; Length 4233;  
Best local Similarity 63.9%; Pred. No. 5,6e-163;  
Matches 895; Conservative 0; Mismatches 496; Indels 9; Gaps 2;

```

OY 162 cagggtgaccctgtatattgttgaatagtgcttgccttgatttttggttacataca 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 355 CATGTAGCCCTACTATTATACAGGGGCGATTGTGCGGTGCTCATCTGTTGCCATACATCCA 414
OY 222 gatttccttgtagtataactgcagcacagacccaagaagatctcgaanaacagtttt 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 415 GGTTCACATTTTGGTCCCTGGAGCGTGGGAGCAAAATATACAAAGATTAGGAGAACTTTT 474
OY 282 tcaatcagtttggcacagagacatcgctggtttgtagtcgtgacatcggtgtacaa 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 475 CCATGCCATCATGATGATGAGGATGAGCTGTGACATGATGACGCTGGGGACCTCAA 534
OY 342 cactcgcatgaca--gacatggacaatacagtagtattgtagtaagaatggctct 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 535 CACCGCGCTCACAGATGACGCTCTCCAAATTAATGACGAATGGTGTGACAAACTTGGAAT 594
OY 399 gttgttcaaaaacatgctacttttgcattggcctggcagttggttttggtagaaggtc 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 595 GTTCTTTCAGTCCATTAAGACATTTTTCAGCCGGTTTATTAATAGGATTTTATAGTGTTG 654
OY 459 gaaactcaacctagtgactatccacgctccctcttataatgcttcaagcgacgatg 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 655 GAAGCTAACCTTGTAAATTTTGGCCGCTCATGCTTATTTGGGTGTCATCTGCCATGTG 714
OY 519 ttctagagatgcatcatcatgaccagtaagaatlaagtgccatcttccaaagctgggac 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 715 GGCAAAAGTACTGACTTATTATTACTTAATTAAGAACTCCAGGCTTATGCGAAAGCTGGAGC 774
OY 579 tgtggcagaagaagctctgtcatcaatccgaacagtcataagcctttagggcccaagaa 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 775 AGTTCGCCAGAGAGCTTAGCAGCCATGAGAACTGTGATTTGGTTGGAGACAAAGAA 834
OY 639 agaacttcaaaagttatcacagaatctcaagaatgtaagaatttggcatabaaagac 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 835 GGAATCTTAAAGGTACATATAAAATTTAGAAAGAGCTAAAGAGTTGGCATTAAGAAAGC 894
OY 699 tataccttcaaaagctctctgtgctgctgtgactctttagaattggaactagagact 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 895 CATCAGGCGCAACATTTCCATAGTATGCTTACCTGTTGCTATGCTCTTATAGCACT 954
OY 759 tgccttttggtagtagaacctccttgacttcttaatgagaacctgtagatbaccatcgagac 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 955 GGCATTTCGTGTAATGGAGACTCCTTGCTGCTCT--CTCAATTAATTAATTTATTTAGACA 1008
OY 819 tgtcttctgcttttcttcttagtgtaataccatagcagttatgcatctggagcgccgtccc 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1009 AGTGCTTACCGCTCTTCTCTATTATTTATTTGGGACTTTCAGTAAATTTGACATTTAGCCCC 1068
OY 879 tcaacttgaacctcgcgaatagcccgagagagctgccttcatatttccaggttatiga 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1069 AAACATTAAGAGCTTTCGCAAAATGCAAGAGGGGACGCTATGAATCTTCAAGATATTGCA 1128
OY 939 taagaanaaccagtagatagaacttttccacagctgtagatataaactgtaatacagaag 998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1129 TAATAGGCAAGCATTTGACAGCTTCTCAACCAAGGAGCAACCAAGAGATATATAGG 1188
OY 999 aacgttgaatttaaaaatgttcttcaatataccatcaagaacatcatacattct 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1189 AAATTTGAAATTTAAATTTTACTTCAATACCAATCACAAAGTGAAGTTAAATCTT 1248
OY 1059 gaaaggtcgaactcagaatlaagttctggagagacagtcgctgtgctgctcgaatg 1118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1249 GAAGGCGCTCAACCTGAAGGTGAAGAGCGGGAGACGGTGAAGCTTGTTGGCAACAGTGG 1308
OY 1119 cagttggaagagtagcgttagtccagctcttcgagaggtatatagtaccgagatgagctt 1178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1309 CTGTGGGAAAGAACACAACTGCTCAGCTGCTGAGAGGCTCTACGACCCCATAGAGGGCGCA 1368
OY 1179 tateatgtgtgtagaagatgacatagacatttaaatgtggcaattatcgagacacat 1238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1369 GGTCTAGTATGACGAGACAGGACATCAGCAATGATGAGATCTGCGGGAATCAT 1428

```

```

OY 1239 tggagtggttagcaagaagccgtttgttcgggaccacacatcaagtaataatcaagta 1298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1429 TGGGGTGGTAGATCAGAAACCCGCTGTTTCCACACAGAAATGGCGAAACATTCGCTCA 1488
OY 1299 tggacagatgagtgtgactgtagaagatgtaggaagcagaaggaagcaaatgctga 1358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1489 TGGCGGCAAAAACGTCCACCATGATGATGAGAAACCTGTCAAGGAGCAAGCAAGCCCTA 1548
OY 1359 tgatttcaatcagaggttccataaatttaattacatttggtagggggaagaaagagctca 1418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1549 TGACTTTCATCATGAAATCTGCCACCAAAATTAACACCTTGTTGGTGTGAGACAGGGGCGCA 1608
OY 1419 aatgagtgaagggcagaacaagaagatcgcaaltgctgctgcttagtgcgaacccaa 1478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1609 GGTAGTGGGGGACAGAAACAGAGATGCGCATTTGCCGGGCCCTGATGCCGACCCCA 1668
OY 1479 gattcgtatttagtagtaggtagtgcgtcccttgatttcaagaagaagtcagctgttca 1538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1659 GATCCTTTGTGTGATGAGGCGCACGCTCAGCCTTGACACAGAAAGCAAGCCGTGTCA 1728
OY 1539 agtgcactggagaagata 1558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1729 GCGCGCTCTGATTAAGCTTA 1748

```

```

RESULT 10
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Field
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; EARLIER FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

```

```

Query Match 29.3%; Score 568.4; DB 4; Length 4233;
Best Local Similarity 63.9%; Pred. No. 5.6e-163;
Matches 895; Conservative 0; Mismatches 496; Indels 9; Gaps 2;

```

```

OY 162 cagggtgaccctgtatattgttgaatagtgcttgccttgatttttggttacataca 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 355 CATGTAGCCCTACTATTATACAGGGGCGATTGTGCGGTGCTCATCTGTTGCCATACATCCA 414
OY 222 gatttccttgtagtataactgcagcacagacccaagaagatctcgaanaacagtttt 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 415 GGTTCACATTTTGGTCCCTGGAGCGTGGGAGCAAAATATACAAAGATTAGGAGAACTTTT 474
OY 282 tcaatcagtttggcacagagacatcgctggtttgtagtcgtgacatcggtgtacaa 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 475 CCATGCCATCATGATGATGAGGATGAGCTGTGACATGATGACGCTGGGGACCTCAA 534
OY 342 cactcgcatgaca--gacatggacaatacagtagtattgtagtaagaatggctct 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 535 CACCGCGCTCACAGATGACGCTCTCCAAATTAATGACGAATGGTGTGACAAACTTGGAAT 594
OY 399 gttgttcaaaaacatgctacttttgcattggcctggcagttggttttggtagaaggtc 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```







Oy	172	ctgattatgttggaagaagtgctgtgccttgatttttggttaatacagatttcttg	231
Db	457	cttgacttctgataaccctggatctggcaggttgctacgtctgataagtaagtaagctggc	516
Oy	232	tgaattataactcagacagacagacagagagatttcgaanaacagtttttccattcagtt	291
Db	517	ttcatctatattaccggaaacacggccacgcaggaagatccggacgtatttaccttgagctatcc	576
Oy	292	ttggcaagaagacatcggctggttttgatagctgtgacatcgtgtgaactaaactcgcagt	351
Db	577	ctggccaggaacatctggctcttttttgatatactgcgtgcggggaaagtacacaccctgtata	636
Oy	352	acagacattgacaaa---atcagttgtgttatgtgaagaagaattgttcctgtgttccaa	408
Db	637	acaccccatatmaaaccttatccagagatggcattttcggaagaagtcggtctacatttgact	696
Oy	409	aacatgtcactcttccgatttgcattgctcgtgcagttggttttgaaggcttgaactcc	468
Db	697	gcccctggcgcacatttcgtaacagcatttatttcgcctacgtacataactggaaagttggct	756
Oy	469	cttagtgactctatcaacgttctcctcttataatggtctcagcgcagcatggttctagatg	528
Db	757	ctaatattggcagctcaaacataattgtggcccttcgttcacacatgggggtgtgtctcagatt	816
Oy	529	gtcatctcatcttgaaccagtaagaagattaaagtccattccataacagctcgtgtgcaaa	588
Db	817	atcatcaagatmaacgcaaaaagtcgcttgacacgtacagctgcagggccacgtcttgcgaa	876
Oy	589	gaagctcttgatcaatccagacagtcatagcctttagggccaggaagaagaactcaa	648
Db	877	gagctcatcagctccatcaacaaaatggccacaggtgtggcaccacaaagacttgcgaag	936
Oy	649	aggtatcacagaatctcacaagaatgcaagaagatttggcagataaagaagactatgcttca	708
Db	937	cnagrttagctccactttagacgaagctggaataatggggacaacaaaccagatttgcattg	996
Oy	709	aaagtgtcctcttggtgctgtgtaactcttataatgaatgaacattatgactgtctttgg	768
Db	997	ggtttcatgatgtggccgcacattgttgaccttattgtactcmaactacgcttggcttgg	1056
Oy	769	tatgaaaccctccttgatcttctaataatgaaacctgtgataacatcggagactgttctgtc	828
Db	1057	atggcttctcgtgttctcctgtagatgg-----tgacgtcgaatgtgggaatatttcttca	1110
Oy	829	gtttctcttagtgaatccatacgaatgaatttcgaatgagcagcagctccctccactttaa	888
Db	1111	gtttcctcagtgccatcttgatcggatgcttctccttggggaacgattagccaatgctcaa	1170
Oy	889	accttcgcaatagcccgagagctgcctcttcataatttccaggtattatgataagaacc	948
Db	1171	gcatttttcaaacgcgtgtggccggccggccgaagaatatttggaaacatgatgcgcacgtcc	1230
Oy	949	agataagataaactcttccacagctgtgataataaccgaaatccataagaagaaactgtgaa	1008
Db	1231	ccatttaatatctcatttttgcgaacgaaggaaagacgttcgacatttttgaggccacattgag	1290
Oy	1009	tttaaaatglttcttcttcaattatccatacaagaacctatcataagattctbtaaaagtltc	1068
Db	1291	tttagcagatgtcaagcattatttttccactctgacccgagccagctcacgcgtatggagatgtt	1350
Oy	1069	aattcgaattaaagtcttgtagaagaacagtcgccttgtgtgtgtctcaatgycagfvggaa	1128
Db	1351	ttctctgtcaantgcccgcttgaaaaaacacccgctttagtgcgcctctgcctcttgaaaa	1410
Oy	1129	agtcaggttagtccagctctcgcagaggttatatgatccggaatgcatgtgcttatacagttg	1188
Db	1411	agtagcgggtgctggcttggttgagcattctacatgcccattgctggcgatagcgttttgcgtg	1470
Oy	1189	gatagaatgacatcagagactttaaatgtcggcattatcgagaccataatgagatggtc	1248
Db	1471	gatggccatgcatacgaagcctcattctccgtctggcttccgcaacagatcttcttggctt	1530

[illegible]

```

RESULT 13
US-08-996-545-3
: Sequence 3, Application US/08996545
: Patent No. 5928898
: GENERAL INFORMATION:
: APPLICANT: Skatrud, Paul L.
: APPLICANT: de Maard, Maarten A.
: APPLICANT: Peery, Robert B.
: APPLICANT: Andrade, Alan C.
: TITLE OF INVENTION: Multiple Drug Resistance Gene atpD of
: TITLE OF INVENTION: Aspergillus nidulans
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: U.S.
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/996,545
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Webster, Thomas D.
: REGISTRATION NUMBER: 39,872
: REFERENCE/DOCKET NUMBER: X-11766
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-3334
: TELEFAX: 317-276-2763
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4002 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
US-08-996-545-3

```



Db	1351	TCCTCTCATATCCCGCTGCAAAACAAACCGCTTATGTGCGCCCTCTGCTCTGAAAA	1410
Qy	1129	aataagtgatgcacgctctcgcagaagttataagatccgagatgagcttatacagtg	1188
Db	1411	AGTACGCTGTGGCGTGGTGTAGCGATTCACATGCTGTTCGGGTACGGTTTTC	1470
Qy	1189	gatagagaataactacagagctttaaattgctgcgcatatcgcagacataatggagtgct	1248
Db	1471	GATGACCATATGATCAAGAGACCTCATATCTCGGCTTGCCACACAGATCTCTTGGTT	1530
Qy	1249	agtcagaagacctgtttgttctggagccaccatcagtaacaata	1292
Db	1531	AGCCAGGAGCGCTGTTCTTTTGGCAGAGATTTTAAAGATTTATAGCACGGCTTCATC	1590
Qy	1293	-----caagtatgacgaagtgtgtgactgtatgaagaatgagatgagaagaagca	1341
Db	1591	GGCACAAATGACGAAGAAATCCGAGATNAAGTTCGGGAATCATGAGAACGGGGCA	1650
Qy	1342	agggagaagcaatgcgctatgatttatacagagcttcccaataatataatacattgta	1401
Db	1651	AAAAATGGCGAATCTCAGATCACTTATTTACTCTCTGCGTGAAGTTATGACACAAATGTT	1710
Qy	1402	ggggaaaaaagagactcaaatgagtgaggagcagaacaaagagatgcgaattgctgtgc	1461
Db	1711	GGGCGAGCTGCTTTTCCCTTTCAGTGGCCGCAAAACAGCGATTCATATCCCGCTCC	1770
Qy	1462	ttagcttgaaaccccaagatctcgattttagatgagctatcgtctccctgattcaga	1521
Db	1771	GTTGTATATGACCAAAAATCTCTGCTCGATGTAAGATTACTTGCGCTTGACACAAA	1830
Qy	1522	aagcaatgcagctgttcaagctgcacgcaggaag	1555
Db	1831	TTCGAGAGCGTGTCTCAACGACGCTTTGAGAGGG	1864

TELEPHONE: 317-276-3334  
TELEFAX: 317-276-2763

TELECOMMUNICATION INFORMATION  
TELEPHONE: 317-276-3334

TELEFAX: 317-276-2763

us-09-873-409-16.rni

1 INFORMATION FOR SEQ ID NO: 3  
2 SEQUENCE CHARACTERISTICS:  
3 LENGTH: 4002 base pairs  
4 TYPE: nucleic acid  
5 STRANDEDNESS: single  
6 TOPOLOGY: linear  
7 MOLECULE TYPE: mRNA  
8 HYPOTHEETICAL: NO  
9 ANTI-SENSE: NO  
US-09-328-320-3

172	ctgtattatglttgaacagtgctgttcctacactttat+++	...cnaachnes 668; Indels 36; Gaps 3,
OY		

```

Db      1291  UUACGCCAAGUCACGAUAAUUUACCACACUAGACCCAGAGUCACACCGUCAAUUGAGAGAUUGU 135
QY      1069  aatcccaagaattaaigtctggagaaagacagtcgttgctgctcgttcaatgacgtgaggag 1128
Db      1351  UCUCUGUCAAUGCCCGCGGAAAAACCAACGCGUUUAGUCGGCCCUUGGCUUGGAAAA 1410
QY      1129  agtcaagtagtccagcgtctctgcagaggtatataatgatccgagatgagcttatacatgtg 1188
Db      1411  AGUACGGUGGUCGCGCTUUGGUGAGCGAAUUCUACAUGCCGUCUCCGGAACGCUUUGCUG 1470
QY      1189  gatgagatgacatcagagcctttaaattgycggcattatcgagacacatattgagtggt 1248
Db      1471  GAUUGGCCAGUAGACUACAAGACCCUCAAUUCGCCUGCGCUUGCCCAACAGAUUCUUUGGU 1530
QY      1249  agtcaagagcctgtttctgttcggagccaccatcagtaacaatat----- 1292
Db      1531  AGCCAGAGACCCUGUUCUUUUGGACACAGAUUUUAAGAUAUUUAGCACCGUCUCUAC 1590
QY      1293  -----caagtatgacagagatgatgtgacgtatgaaagatgagagagagca 1341
Db      1591  GGCACAAAGUACGAGAAUGAAUUCGAGGAUAAAGUUCGCGGAAUCUACAAGAACGCGCA 1650
QY      1342  agggaaagcaatgctgattatattcatcagaggtctccctaataatcatatgtgta 1401
Db      1651  AAAAUGCGAAUUCUACAACUUAUUUACUCCUGCCUGAAGUUUAGAACCAAUUGU 1710
QY      1402  ggggaaaaagagctcaaatgagtggagggagaaaaacagagatcgaattgctgtgcc 1461
Db      1711  GGGCACGUGGCUUUCUUCUUCAGUGGCGCAAGAAACAGCGCAUUGCAAUCCGCGUGCC 1770
QY      1462  ttagtctgaaaccccaagatcttgattttagatgaggtcagtcgtcgttgatcgaa 1521
Db      1771  GUUGUAGUGACGAAAAAUUCUUCUCCUGGAUAGUACUACUUGGCUUGGACACAAAA 1830
QY      1522  agcaagtcagctgttcaagatcgtgcactgaggaag 1555
Db      1831  UCCGAAAGCCUGUGUCAAGCAACGUUGGAGAGG 1864

```

2002

us-09-873-409-16.rni

**THIS PAGE BLANK (USPTO)**

...  
Tue Apr 23 09:35:22